PERCEPTIONS OF AIR FORCE BASE-LEVEL TRANSPORTATION OFFICERS TOWARDS THE EFFECTIVENESS OF AIR FORCE BASE-LEVEL TRANSPORTATION PERFORMANCE MEASUREMENT INDICATORS

THESIS

Kevin N. Brewer, 1st Lt, USAF

AFIT/GLM/LSM/89S-3

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

2

AFIT/GLM/LSM/89S-3

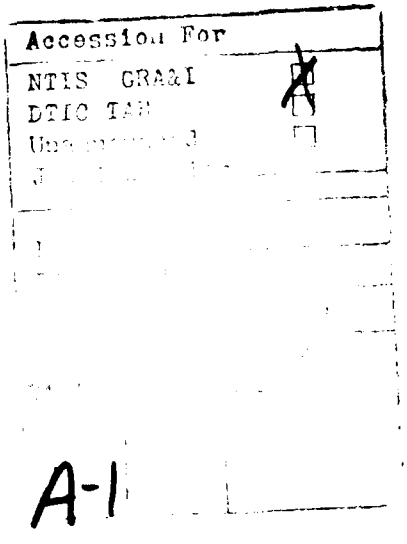
PERCEPTIONS OF AIR FORCE BASE-LEVEL TRANSPORTATION OFFICERS TOWARDS THE EFFECTIVENESS OF AIR FORCE BASE-LEVEL TRANSPORTATION PERFORMANCE MEASUREMENT INDICATORS

THESIS

Kevin N. Brewer, 1st Lt, USAF

AFIT/GLM/LSM/89S-3

The contents of the document are technically accurate, and no sensitive items, detrimental ideas, or deleterious information is contained therein. Furthermore, the views expressed in the document are those of the author and do not necessarily reflect the views of the School of Systems and Logistics, the Air University, the United States Air Force, or the Department of Defense.



AFIT/GLM/LSM/89S-3

PERCEPTIONS OF AIR FORCE BASE-LEVEL TRANSPORTATION OFFICERS TOWARDS THE EFFECTIVENESS OF AIR FORCE BASE-LEVEL TRANSPORTATION PERFORMANCE MEASUREMENT INDICATORS

THESIS

Presented to the Faculty of the School of Logistics
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Logistics Management

Kevin N. Brewer, M.S.
First Lieutenant, USAF

September 1989

Acknowledgements

I wish to thank my advisor, Major Robert McCauley, for his support and assistance in the preparation of this thesis. I would also like to thank my classmates for their understanding and patience during trying times.

Table of Contents

LIST OF TABLES

List of Figures

Abstract

The purpose of this study was to examine the perceptions of selected base-level transportation performance measurement indicators. These included the Vehicle Integrated Management System (VIMS), as well as certain vehicle operations indicators. Transportation squadron commanders and certain staff officers within the continental United States belonging to the Strategic Air Command (SAC), Tactical Air Command (TAC), and Military Airlift Command (MAC) were surveyed. The study examined the officers' perceptions against seven criteria for good performance measurements developed by the A.T. Kearney consultant firm.

1. Validity
2. Coverage
3. Comparability
4. Completeness
5. Usefulness
6. Compatibility
7. Cost Effectiveness

An examination of the perceptual surveys found that SAC personnel have a slightly more favorable perception that the performance measurements met the criteria than the other commands, followed by TAC, and then MAC. A comparison between Line and Staff officers revealed Line officers have a slightly more favorable perception that the performance measurements met the criteria.

Analysis of the rankings of the performance measurements indicated the five most important to the respondents were

1. Percentage Hours VDM-Registered
2. Percentage Hours VDP-Registered
3. Percentage Hours VOC-Registered
4. Safety Inspection Overdue-Registered
5. Scheduled Inspection Overdue-Registered

One of the recommendations provided to continue the study based on a Delphi approach using other military service transportation personnel, members of the academic transportation community, and commercial transportation personnel.

PERCEPTIONS OF AIR FORCE BASE-LEVEL TRANSPORTATION OFFICERS TOWARDS THE EFFECTIVENESS OF AIR FORCE BASE-LEVEL TRANSPORTATION PERFORMANCE MEASUREMENT INDICATORS

I. Introduction

Overview

Performance measurement has become increasingly important to today's managers. Whether in the civilian or military context, performance measurement can be directly translated into dollars. In the case of national defense, dollars contribute to combat capability.

The United States spent an estimated $713.2 billion dollars on transportation in 1984 (43:11), or 19.47% of the Gross National Product. The significant expenditure of national resources in the transportation sector points out the potential associated with effective productivity and performance measurement systems. A mere one percent increase in transportation productivity in the United States would result in cost savings of seven billion dollars.

The organization of this chapter is as follows. First a background of the present productivity situation is presented, beginning at the national level and narrowing to the major command level. The general issue at base-level transportation squadrons, including the overall management

question, is examined in the second section. The specific problem statement and its importance follows in the third section. Fourth, the investigative questions of this research are established. The definition of performance and performance measurement indicator as well as what should be included in this definition are stated in the fifth section. Sixth, a brief synopsis of the state of productivity in the United States is presented. Seventh, additional definitions of performance and productivity are given. In the eighth section, an analysis of the state of performance or productivity measurement in the Air Force is presented. Finally, the scope, limitations, and assumptions of this paper are given.

Background

Performance is "presumed synonymous with productivity; in fact only a subset of performance variables can be equated with productivity."(7:19) The effectiveness or efficiency of the organization often is equated with its performance. In the strictest sense, productivity refers to performance efficiency, while performance measurement addresses either organizational efficiency or effectiveness (7:19). However, an effective performance measurement system is essential to the development and execution of a sound program of productivity enhancement.

President Jimmy Carter recognized the need for productivity improvement when in 1978 he established the

National Productivity Council. The council's charter was to coordinate all United States productivity improvement achievements (2:3). On the subject of productivity, President Carter stated

> I have established this council in recognition of the vital role productivity plays in the nation's economy by helping control inflation, making the U.S. goods more competitive in world markets, and increasing the real income of the American worker . . . improved productivity is vital to the social and economic well-being of our nation. The Federal Government can make a major contribution to improving productivity. I expect all agencies to cooperate and assist the council in meeting its responsibilities so we realize maximum benefit from the Federal effort to improve productivity growth. (45:3)

The Department of Defense recognized the need for increased productivity in the military when it established the DOD Productivity Improvement Program under the guidance of Department of Defense Directive (DODD) 5010.31, in 1979. The goal of this program was to increase productivity in the DOD (36:1). A fundamental part of this program was the development of an accurate performance measurement system. Howell and Van Sickle described the intent of the DOD Productivity Improvement Program as follows.

> DoD Directive 5010.31 establishes the policy of focusing management attention on the achievement of maximum defense output within available resource levels by seeking out and exploiting opportunities for improved methods of operations in consonance with the defense preparedness mission. The directive further states that, productivity measurement, enhancement and evaluation will be an integral element of resource management . . . The Directive prescribes a labor-oriented mode, but allows for total product or unit cost measures if available. (31:6)

The Air Force developed a comprehensive productivity improvement plan that directed all major commands to formulate their own productivity plans. Each major command appointed "productivity principals" as the individuals responsible for coordinating and reporting productivity accomplishments to United States Air Force Headquarters (2:4). Following issuance of the plan, General James A. Hill, Vice Chief of Staff, stated:

> Productivity has received increased emphasis at all levels of government, and had consistently surfaced as key factor during congressional debates on Military Appropriations. If we are to continue obtaining the necessary funding for vital Air Force Programs everyone at all organizational levels must actively seek more productive means of accomplishing their jobs. We therefore urge your full support for this plan. (45:4)

The potential impact of productivity on the Department of Defense is illustrated by General Louis L. Wilson, Jr.'s statement,

> The Air Force is facing one of the most austere times in its history. In spite of increased defense budgets, our buying power has eroded with the net result that we have to do more with less. To meet this challenge, we need to fully utilize our most costly and important resource . . . people . . . by instilling in them a sense of urgency about their important role in the conduct of the Nation's critical enterprise . . . national security . . . and in doing so we must increase their productivity. (2:2)

The following example illustrates the potential benefits to base-level transportation operations and consistent cost savings derived from an effective performance measurement system. Minot Air Force Base manages a vehicle fleet valued

at over $20 million. Each vehicle must be justified based on its intended usage. Vehicle usage is measured by the base-level vehicle operations function. When a vehicle is brought in for maintenance, usage of other vehicles increases. The longer a vehicle remains out of commission, the more vehicles are needed in the field to compensate its temporary loss. When dealing in terms of $20 million dollars for just one base, the cost of compensating vehicles is high.

The length of time the vehicle remains "out of commission" is measured by the performance measurement indicator, Vehicle Out of Commission (VOC) rate, found in the Vehicle Information Measurement System (VIMS) (17:1). Unfortunately, the relative priority or importance to the mission is not considered by this performance measurement indicator. It is conceivable an acceptable VOC rate may be attained, and yet, the mission needs not be met.

Since each MAJCOM and transportation squadron should have a productivity program, one of the functions of this program should be the periodic evaluation of the performance measurement systems. The performance measurement indicators themselves should be part of this evaluation.

Many times the performance measurement system is updated through automation, but the performance measurement indicators are not updated to reflect system changes. These changes are wrought by modernization of equipment in the field, data automation, and the civilianization of the Air Force workforce.

General Issue

The base-level transportation squadron has unique performance measurement problems complicating productivity assessment of the unit. A part of the DOD, the base-level transportation squadron is a nonprofit organization. A transportation squadron is both a production and service organization. With diverse requirements placed on transportation squadrons and managers, accurate indicators of performance are essential.

The vehicle maintenance function is a job shop that repairs or modifies equipment from vehicles to lawn mowers. This branch tracks vehicle maintenance records, determines useable vehicle lifespans, and schedules preventive maintenance actions. The vehicle operations branch provides a fleet management, customer service and vehicle tracking function. Every base vehicle must be obtained through, and accounted for by this office. This branch also provides military driver's licensing and training for other base organizations on special purpose vehicles such as semitractors. The traffic management branch has dual functions of personnel and personal property transport as well as the packing, crating, shipping and receiving of munitions and equipment. Because of the broad and diverse nature of these functions, it is very difficult to judge performance in transportation functions without effective performance measurement indicators.

Major Jont Johnson, Chief of Transportation, Wright-Patterson Air Force Base, Ohio captured the difficulties associated with transportation squadron performance measurement when he posed the question, "How do you measure customer service? How do you quantify customer service?" (32)

When inundated with squadron performance data Lt. Col. Dan E. King, the 1988 Transportation Officer of the Year for the Strategic Air Command, and the Transportation Commander at Minot Air Force Base, North Dakota said,

> There is so much performance data, a commander cannot look at all of it, all of the time. I need to know what information is important to me in order to manage my squadron. (35)

The management question at hand is, "Do Air Force transportation squadrons have effective performance measurement indicators?"

There are many other questions that need to be answered to respond to the overall management question. First, an analysis must be made to determine the adequacy of present performance measurement indicators in meeting stringent criteria for effective performance measurement systems.

Specific Problem Statement.

The experts in Air Force transportation performance measurement are people regularly using transportation measurement indicators. Specifically, the experts are the base-level transportation squadron commanders and their branch chiefs. Their perceptions can assist in determining whether the Air Force has effective indicators of transportation

squadron success. The research question addressed here is "What are the perceptions of Air Force transportation officers toward the effectiveness of transportation base-level performance measurement indicators in reflecting squadron performance?"

Investigative Questions

There are three primary investigative questions that must be answered.

1. What are the base-level transportation performance measurement indicators?
2. What are the essential characteristics of performance measurement indicators?
3. What are the actual perceptions of transportation officers towards the base-level performance measurements using a reference standard of seven currently accepted criteria.?

The first investigative question is an important one to answer, because not all data available to a transportation squadron are performance related. Much of the data available in a transportation squadron is related to the squadron activity levels.

The answer to investigative question two is equally important. How the performance measurements are derived provides insight to its intended use. For example, if the measurement is output per man-hour, then exactly where the number for the man-hour factor comes from is important. If the origin of the factor was not known, the factor could be misinterpreted or "adjusted" by the transportation manager.

The users perceptions (captured by investigative question three) are essential in determining the effectiveness of the measurement indicators. If the people using the measurement indicators do not find them effective or useful, then the cost and effort of collecting, recording, and analyzing the measurement indicators is wasted. If this is the case, the performance measurement indicators are useless and should no longer be collected.

Definition of Performance

Performance has been defined in many ways. A general definition of performance is actual output over expected output. However, performance involves, more then output (34:237). If performance considered outputs without considering amount of input, any amount of resources could be used to achieve the desired level of performance with no deleterious effects.

A true performance definition should include the relationship between performance and utilization (input available over input required) as well as the relationship between performance and productivity (34:237). It would also include the relationship between performance and quality.

The definition of performance should include the relationship between operating technology and its level of inherent productivity (34:38). For example, assume the vehicle maintenance function obtains a new computer analyzer that speeds the repair process with no other increase in

resources. The performance measurement indicator must consider this factor when comparisons are made between historical and current levels, and between organizations.

In this paper, a performance measurement indicator is defined as "the relative element used to evaluate macro-, micro-, long-term, short-term, flow, static, functional, and overall performance by evaluating the inputs, outputs, transformations (including level of technology), and productivity in a manufacturing or nonmanufacturing operation" (10:229). Other definitions concerning productivity and performance measurement can be found in Appendix A.

As stated above, productivity and performance frequently are used synonymously. Because of this, productivity and performance are used interchangeably within this paper. One source states although some researchers differentiate between these terms, states many believe they are similar enough or the area is grey enough to permit their interchangeable use (7:19).

U.S. Productivity

Domestically, increases in U.S. productivity, generally measured as real gross national product per worker has been slowing in rate since 1970 (34:1). The decrease in the rate of growth can be attributed to many factors. Many of these factors are related directly to performance levels of companies. One factor is the increase in energy cost. This results in the search for alternative production techniques

which may not result in an increase in output but do represent a considerable cost (24:320).

Another factor is the shift from agricultural and manufacturing emphases to the service and government sectors. Not only are increases in productivity difficult to realize in the service and government sectors, productivity is harder to measure in these areas. The number of cars produced per day is a more concrete and precise measure of productivity in the automotive sector than the number of clients serviced in the service sector. The former is a discrete measurement. The latter involves qualitative variables, such as the inability to provide for customers' needs. Other qualitative variables such as, the size and complexity of clients' needs, and service times may vary greatly. The production of the car, however, would follow a relatively standard amount of time.

Another factor is the inverse relationship between productivity and product variety, customization, and new product introduction. To market products properly and compete according to these three tenets, a price must be paid, and often that price is a reduction in productivity because of the division in resources involved (39:320).

Other Definitions of Productivity

There are many measures of productivity. The standard definition of productivity is the ratio of input to output (34:237). There are two general approaches to productivity measurement. The first is through the financial accounting

system and is called the "Economist's Approach" (24:245). This uses dollars as a standard unit for measurement data and is used in total industry statistics (2:45). The second is called the Engineer's Approach and removes the dollar from the system and deals only in physical units or data and is used for individual operations (2:45).

Economist's Approach

$$\text{Productivity} = \frac{\text{Output}}{\text{Input}}$$ (Both units measured in terms involving dollars)

Engineer's Approach

$$\text{Productivity} = \frac{\text{Useful Output (Work)}}{\text{Input (Energy)}}$$

(2:20)

A comparison of the two definitions shows the various peculiarities of each. The Economist's Approach requires a performance measure exceeding a value of one for the process to add value to the input (2:20). The Engineer's Approach requires an upper limit of one as more work cannot be done than energy used (2:20).

Productivity sometimes is defined as a specific relationship. One generally used relationship is one of labor efficiency which is standard hours/actual hours. Another relationship is labor effectiveness. Labor efficiency and effectiveness alone cannot measure productivity because these measures do not include capital or materials. However, labor efficiency and labor effectiveness coupled with other current measures such as inventory turn ratio, value per employee,

ratios of direct versus indirect labor and gross margins relate a generally accurate measure of productivity (39:716-717).

In recent years, productivity has been studied extensively. Many theories have been proposed to explain why the U.S. productivity rate of increase has dropped from 3% during the early sixties to 2% by 1985 (28:320).

Productivity in the U.S., based on a Bureau of Labor Statistics bilateral productivity index of 100 in 1957, rose to just under 200 in 1982 while Japan rose to 750, France and Germany rose to 350, and Canada and the United Kingdom rose to 250 (34:1)(7:42). However, all is not doom and gloom in the U.S. production picture. The productivity levels of other major industrial countries have not risen to the level of the United States. Nonetheless, other countries are closing the gap. On an index of gross domestic product per employed person, using the United States as a benchmark of 100, France was at 95 followed by Canada with 94 and Germany with 92. Far back on the list was Japan with 73 (34:14). In nominal terms, the U.S. is still far ahead in overall productivity.

In 1978, only 30% of U.S. companies had started any productivity measurement programs (34:47). Five years later, half were still in the starting phase (34:47). Because of the tremendous lead times involved in establishing such programs, gathering data, interpreting results, making any necessary changes, gathering the resulting data, and determining what

the data means; the net results of these programs will not be seen until the next decade and beyond.

Scope/Limitations

The results of this study will formulate an answer to the overall management question, "Do Air Force transportation squadrons have effective performance measurement indicators?" This research study will survey the target population for their perceptions of selected performance measurement indicators. This research study will not attempt to evaluate every performance measurement indicator used by the base-level transportation squadron but will concentrate on the primary indicators within two branches, vehicle operations and the vehicle maintenance. This study will further limit the measures evaluated. Only performance measurement indicators identified in key interviews with transportation experts and publications will be evaluated.

Summary

The measurement of performance is essential to the continued progress of all levels in the United States to include the U.S. Air Force base-level transportation squadron. This research examines one aspect of the state of performance measurements within the base-level transportation squadron, the perceptions of the officers using these measurements.

Chapter II examines the current literature on the subject of performance measurement within the civilian sector as well as the Department of Defense (DOD). This review includes a

survey of current Department of Defense, and Air Force transportation regulations dealing with productivity.

Chapter III describes the methodology used to accomplish this study and justifies why this method was chosen. The chapter delineates the sample and sampling plan. Statistical tests applied to the survey data are defined, explained, and justified.

Chapter IV presents the results of the study. Tables of important demographic data; percentage descriptive statistics; relative ranking of transportation performance measurement indicators; frequencies of response; and the results of chosen statistical tests are contained within this chapter. Chapter V analyzes the data obtained and presented in Chapter IV. Chapter VI, the final chapter, sites the conclusions of this study and its specific recommendations. This chapter includes a general outline for subsequent studies of performance measurement within the Air Force transportation community.

II. Literature Review

Introduction

Increasing performance or productivity (used here interchangeably) is an area of great focus during these years of tight budgets. An attempt was made to find some type of model or framework within the field to use as a guide but it was soon evident that no prevailing model exists. In fact, it was stated:

> research on productivity is segmented according to academic discipline (such as psychology or economics) and that each discipline pursues its own level of analysis. Forcing an integration of the two is therefore unnecessary and probably impossible . . . When it comes to productivity research and theory, let a hundred flowers bloom . . . It is perplexing and inexplicable that industrial and organizational psychology has devoted so little effort to the systematic explication of the performance domain . . . The absence of models and theories of performance is a gap that needs to be closed. (7:77,419)

In light of the fact no definitive models exist, literature was reviewed for those areas pertinent to this particular study. This chapter is presented in four major sections.

In the first section, special attention is directed towards those works which examined performance measurement criteria similar to A.T. Kearney's list of seven performance criteria (34:43). Here the works are classified according to what characteristics the researchers feel are most necessary.

In the second section, special attention is given to those works which examined at the actual system outputs: total

productivity (capital and labor), standard costing, and inventory. These included works which addressed several approaches: (1) cost-based measures (2) transportation-related measures, (3) resource tracking measures, and (4) measures in nonprofit organizations.

The third section examines current Air Force progress in operational and support areas of performance measurement. This section highlights a (Air Force Logistics Command) commander's philosophy and policies. This section also describes new programs such as the READY RESOURCE maintenance program and the Contractor Performance Assessment Reporting System (CPARS).

The final section, future trends, shows a recent tendency in industry for performance measurements to be developed using throughput time as a prime factor. This section also outlines attitudes toward the characteristics necessary in future performance measurement systems.

As a whole this chapter examines current literature on performance measurement within the civilian, Department of Defense (DOD) and Air Force sectors. A review of current Department of Defense regulations dealing with productivity and Air Force transportation regulations is included.

Performance Definition

Performance has been defined in many ways. This research uses Crawford's definition of performance measurement,

the relative element used to evaluate macro-, micro-, long-term, short-term, flow, static, functional, and overall performance by evaluating the inputs, outputs, transformations (including level of technology), and productivity in a manufacturing or nonmanufacturing operation. (10:240)

Characteristics of Good Performance Measurement Indicators

Most researchers agree on the common elements characterizing effective performance measurements. As represented in Table 1 research indicates general agreement on performance measurement criteria similar to A.T. Kearney's list of seven performance criteria (34:43).

Table 1
Performance Measurement Characteristics According to Author

Source	Kearney	AMH	Tuttle	Tyworth et al.
Characteristics				
1. Validity	X	X	X	X
2. Coverage	X	X	X	X
3. Comparability	X	X	X	X
4. Completeness	X		X	X
5. Usefulness	X	X		X
6. Compatibility	X	X	X	
7. Cost effectiveness	X		X	
8. Acceptability			X	

(34:43)

A.T. Kearney.

There have been many civilian works on productivity and performance measurement in current literature. This section begins with a synopsis of the seven general performance measurement criteria formulated by A.T. Kearney in a study sponsored by the National Council of Physical Distribution Management. At that time seven criteria were found by the consulting firm to be the most comprehensive list currently available (34:43).

To have an effective performance measurement indicator, the indicator should satisfy most of the seven general performance criteria in the Kearney study.

1. Validity
2. Coverage
3. Comparability
4. Completeness
5. Usefulness
6. Compatibility
7. Cost effectiveness
(34:43)

First, the measure must be valid. Does it represent real productivity? An example of this is if

> a part of a receiving function in a warehouse is to unload pallets of goods from trailers using a forklift and move them to a staging area, then pallets-loads per hour would be a valid measure of productivity or performance. Cases per hour or pounds per hour would not, (34:42)

be a valid measure because the case per pallet load or the weight in comparison to size, of the case may vary.

The second criterion is coverage. Does the performance measurement indicator measure all aspects of the resource? The better a measure can account for all uses of that resource the better the resource is tracked. For example, if a driver works two types of jobs during the same shift, say washing vehicles and driving, two separate measures may be needed to track performance (34:43).

Third, is the performance measurement indicator comparable to past measurements of the same activity? For example, if two pallets of hazardous cargo per hour in 1985 is the standard, should the same standard apply in 1989? In

1989, automated labelling of this type of cargo is available, making the task easier and possibly more accurate. Things such as product mix, order size, customer location, and throughput volume must be considered. Normally, performance measurement indicators should be able to be reduced to a common denominator or a standard work unit. In the previous example of a driver doing two jobs, the performance measurement indicator would be reduced to a common standard work hour. Then the amount of time spent on each task could be added to represent the work done per day and from one facility to another (34:44).

The fourth criterion is completeness. Does the measurement indicator include all important outputs from producers? For example, if a driver also does extra duty as a dispatcher during busy periods, is it measured? All important inputs and outputs should be tracked. The absence of one (such as cost) could be disastrous. The thoroughness used in tracking important resources such as cost, labor, energy, floor space etc . . . is essential (34:44).

Fifth, can this performance measurement indicator allow managers to make good decisions from the data? In other words, is the performance measure useful? The performance measurement indicator should guide the manager to a good decision. If this is not the case, then a measure of overall productivity in an organization may be of little use to the manager (34:44). Information that cannot be used is worse

than none at all because its collection represents wasted effort on the organization's part (34:44).

The sixth characteristic is compatibility. Does the measurement indicator require new data or can it be used with existing systems and data? The presence of computers at low levels of an organization has made consideration of this factor essential to the effective development of a performance measurement indicator. Manipulation of large amounts of data by computers may require performance measurements be stored in modular form. The development of a new route of information flow, a computer system for example, is expensive. The cost would be prohibitive to develop a new system for every new performance measurement. The ease of implementation also is affected by the compatibility of the measurement indicator (34:44).

Seventh, are the benefits derived from tracking this performance measurement indicator cost effective? As previously stated, a computer system may cost more than the value of the information. If the new measurement requires a worker to spend an excessive amount of time collecting the data, or another employee to track the measurement, the benefits may not outweigh the costs of data collection (34:43).

American Management Handbook.

The American Management Handbook is the second major source and falls primarily into the category found in Table 1

because of the eight criteria shown below. In the American Management Handbook, work measurement is defined as "a method for determining the amount of output actually produced during a set unit of time" (23:10-36). The measurement indicator must:

1. [be able to be] measured meaningfully.
2. be quantifiable.
3. accurately reflect the amount of work being done.
4. be consistent.
5. be for an appropriate unit of time (not too long or short).
6. be appropriate to task being measured.
7. be meaningful to managers.
8. [be able to use] existing records or electronic data

(23:10-36,10-37)

These roughly correspond to A.T. Kearney's seven criteria for performance measurement criteria.

The handbook also states that "outside factors that undermine productivity . . . should be corrected before final standards are set" (23:10-40). These factors included machine failure, employee illness, and material shortages (23:10-40). They inherently would seem to include manpower levels, budgetary and environmental conditions.

Tuttle

Tuttle is listed in Table 1 because his research contained performance measurement indicator characteristics. In a research report on productivity measurement for the Maryland Center For Productivity, and the Air Force, Thomas Tuttle made several observations based on a survey of Chief Executive Officer definitions of productivity. Tuttle found

that ninety percent of the managers would include "quality, effectiveness and efficiency," seventy percent would include some sort of factor for "work stoppages, waste, shrinkage, sabotage, absenteeism and turnover," sixty percent would include customer satisfaction (45:11). This shows a wide variation in managerial perceptions of productivity.

Tuttle used the same criteria for performance measures as the National Council of Physical Distribution Management except for one. Tuttle did not include usability. In its place, he placed acceptability or supportability. Tuttle felt if an organization did not accept and support the measurement then it would not work. This expansion of the definition an important addition to the other seven criteria (45:77-78).

Units of Productivity or Performance Measurements

Another body of research examines the types of performance measurement systems and the units of productivity needed in the systems. The research reviewed is summarized in Table 2.

Lesser and Roller.

Lesser and Roller are found in Table 2 because their research focused on the input/output relationship. Lesser and Roller found that many researchers identified a direct relationship between facility size and physical productivity (25:98). This may indicate that the size of the facility may need to be considered in the construction or at least in the analysis of transportation performance measurement indicators.

Table 2

Output and Performance Measurement Units According to Author

Units	Input/Output	Capital+Labor	Standard Costing	Inventory
1. Lesser/ Roller	X			
2. Deakin/ Seward		X		
3. Conlon/ Townsend			X	
4. Campbell/ Campbell			X	
5. Ballou	X	X	X	
6. Temple/ Barker/ Sloane			X	
7. Coyle/ Bardi/ Cavinato	X		X	
8. Goldratt			X	X
9. Tyworth/ Cavinato/ Langley			X	
10. Anthony			X	
11. DODD 5010	X			
12. Hanley/ Smith		X		
13. Williams	X			
14. Baumgartel/ Johnson	X	X		
15. Kaneda/ Wallett			X	

Deakin and Seward.

Deakin and Seward state in their book that productivity measurement based on labor alone is crude because it does not consider the quality or changes in the composition of the workforce (especially in technical areas) (11:23-25). This may indicate that some measures should allow for the skill composition of the transportation squadron workforce. Deakin and Seward advocate the use of Total Factor Productivity measures (11:111).

$$\text{Total Factor Productivity} = \frac{\text{output}}{\text{labor + capital + resources + miscellaneous}}$$

(36:23)

These measures take into account capital and labor at the same time "to reach a productivity measure of output per unit of combined, labor and capital, factor input" (11:111).

Conlon and Townsend.

Conlon and Townsend identified four phases to aid in productivity measurements clearly placing them in Table 2 with their concentration on standard cost systems (8:63-64).

1. Evaluate Operations-(Capture data already available)
2. Establish Standards-(To develop target achievements (standards) in each functional area (to include overall organization function).
3. Expense Elements-(To break out and separate each functional elements in broad terms based on the unique cost base for that function.
4. Calculations-(Analysis of actual costs from payroll records, expenses etc . . . (8:63-64)

It is apparent these researchers believe standards should be function specific and include functional cost elements.

Campbell and Campbell.

Campbell and Campbell concur with the majority of performance researchers that dollars should be the primary measure of productivity and this places them in Table 2. These researchers state:

> While physical measures of output might be possible for a single product unit, the common measure for multiple product units is some measure of monetary value. (7:23)

The researchers indicate some of the same attitudes toward the use of performance measurements as other researchers reviewed here. Campbell and Campbell feel that performance measurement is specific to individual units for short-run application rather than for comparisons between units (7:25). This feeling is shared by Baumgartel and Johnson (5:71-72).

Ballou.

The concentration on costing systems places Ballou in Table 2. Dr. Ronald Ballou states the "methods that help to measure, compare and guide physical distribution performance have not received, [as much attention as other areas such as computer based information systems]" (4:837). Ballou feels performance measurements should capture trends because the measure of one period by itself reveals little useful information (4:837). Dr. Ballou asserts the measures do not compare input to output, rather they show

> absolute level of either output or input rather than a relationship between the two, which would help show whether resources are being used in an efficient manner and whether inputs are remaining at reasonable levels relative to output. (4:838)

Ballou splits measurements into three major forms:

> Good measurement reporting depends on three major reporting forms: cost-service statement [comparing against time or budget], the productivity report [% of costs], and the performance chart [output vs. cost]. (4:840-841)

From these statements it is clear that Ballou agrees with

other researchers that output should be compared with resources and their costs, i.e. manpower, facilities, and budget.

Temple Barker and Sloane.

Temple, Barker, and Sloane, in a study for the National Council of Physical Distribution Management, ranked productivity assumptions (for transportation in the 1980's) into modal (type of transportation) and three cost components, labor, fuel, and equipment (44:238-243).

Coyle, Bardi and Cavinato.

Coyle, Bardi and Cavinato state "functions costs by driver, vehicle, plant and so on will permit analysis of problem areas [in transportation functions] . . . " (9:406). The authors feel functional costs "collected at the source . . . permit the analysis of individual cost centers . . . " (9:406). A list of 29 sample performance measurement indicators included 16 cost indicators covering a level low enough to include plant, driver, and customer (9:407). This is in agreement with most other research works reviewed that favored the tracking of costs at the lowest level possible.

Goldratt.

According to Goldratt, performance measurements should be based on three categories of measurements (26:4).

1. Throughput: The rate at which the system generates money through sales.
2. Inventory: All the money the system invests in purchasing things the system intends to sell.
3. Operating Expense: All the money the system spends in turning inventory into throughput. (26:4)

Goldratt also believes that all expenses of a subsystem controlled by that subsystem should be assigned to that subsystem (26:16). This would seem to indicate expenses (costs) should be assigned where they occur, at the organization (squadron or branch) level.

Tyworth, Cavinato, and Langley.

Although Tyworth, Cavinato, and Langley's work is referenced in both tables, their research is presented here because the bulk of their work pertained to performance output units and systems. The researchers break the terms productivity, performance, and utilization into three different categories:

> Productivity is the ratio of real output produced to real resource consumed. Utilization is the ratio of used capacity to available capacity. Performance is the ratio of actual output to standard output (or standard hours earned to actual hours). (46:444)

The authors stated the performance measurement indicators must do three things. The indicators must provide information for line managers to make decisions. The indicators must show over time whether the strategies of the organization are effective. The indicators also must inform top management of progress toward the organization's goals (46:157).

Tyworth, Cavinato, and Langley suggested 15 performance measurement indicators for transportation. It is important to note that 10 of the 15 measures involved measurement of the costs involved which places them in Table 2 under standard cost systems (46:446). To be effective, the authors state "these measures need to be broken down by product groups, geographic regions, origin points, modes, and carriers and tracked over time." (46:453) It is clear these researchers feel performance measures cannot be compared across organizations without considering these types of factors during measurement indicator construction or analysis.

Anthony.

Anthony examines the organizational consequences of not having a profit performance measure (1:42-43). The consequences pertaining to this study and the DOD in general are:

1. The presence of an unclear connection between benefits and costs.
2. The difficulty in measuring productivity.
3. The difficulty in comparing the productivity among units in the organization. (1:42-44)

This concentration on the disadvantages of a nonprofit organization places Anthony in Table 2 under standard cost systems. These consequences are of direct relevance to this study and represent much of the problem in examining, constructing, and analyzing transportation performance measurements.

Department of Defense Directive 5010.31.

The Department of Defense recognized the need for increased productivity in the services when it published Department of Defense Directive (DODD) 5010.31, in 1979. The directive's objective was to increase productivity in the DOD (19:1). This was called the DOD Productivity Improvement Program. One aspect of this program was the development of an accurate performance measurement system, The program's purpose was best described by Howell and Van Sickle:

> DoD Directive 5010.31 establishes the policy of focusing management attention on the achievement of maximum defense output within available resource levels by . . . seeking and exploiting opportunities for improved methods of operations in consonance with the defense preparedness mission. The directive further states that, productivity measurement, enhancement and evaluation will be an integral element of resource management . . . The Directive prescribes a labor-oriented mode, but allows for total product or unit cost measures if available. (31:6)

The Air Force developed a comprehensive productivity improvement plan that directed all major commands to make their own productivity plans. Each major command was to appoint "productivity principals" as the individuals responsible for coordinating and reporting productivity accomplishments to United States Air Force Headquarters (2:4).

The definition of productivity, according to the Department of Defense, was a combination of efficiency and effectiveness. Efficiency meant to accomplish the mission correctly at least cost. Effectiveness meant to do those things at the right time. As specifically stated in DOD

guidelines, "The efficiency with which organizations utilize all types of fund resources to accomplish their mission represents total resource productivity" (20:1). A performance index according to DOD guidelines is "the percentage ratio of goods produced or services rendered (outputs) to resources expended (inputs) during a current period in relation to a base period." (20:11)

Hanley and Smith.

In 1976, Hanley and Smith examined "the effect of labor manhour requirement estimates on the measurement of Air Force Civil Engineering (AFCE) productivity" (2:18). This would place Hanley and Smith in Table 2 under "labor+capital." Hanley and Smith found that variation of work estimates significantly differed among civil engineers (29:15). These estimates are important because of their use in computing work standards. The work standards are used to compute a productivity index. The research team concluded, "Comparisons of standard estimates with actual manhour expenditures result in unreliable productivity ratios" (29:15). Presence of such a condition in a customer service/production organization such as civil engineering may indicate existence of a similar problem in other Air Force organizations such as transportation.

Williams.

Dr. Ben Williams is director of the Center for Excellence for Reliability and Maintainability at the Air Force Institute

of Technology, Wright-Patterson Air Force Base, Ohio. Dr. Williams feels the U.S. measures "productivity in the wrong way." (47:9-A) The director feels the resources (inputs) used to meet a specific need must be considered to have an appropriate measure of productivity (47:9-A).

Baumgartel and Johnson.

Baumgartel and Johnson attempted to formulate new performance measurements for a base level civil engineering organization. Their performance measurements took the average values of branch tasks, and divided them by the resources consumed to achieve the level of output (5:24). A series of these measurements over a period of time would show the progress of the organization toward its objectives. The measurements were designed to be used within that organization and not extended to others (5:71-72). The researchers stated that performance measurements must be designed specifically to compare readiness between bases. (5:108). This finding confirms the results of other researchers (33:79-80).

Kaneda and Wallett.

Another civil engineering attempt to develop performance measures was done in 1980 by Kaneda and Wallett. The data was collected from questionnaires completed by military and civilian chiefs of engineering. The research team developed six performance measures from the data (33:79-80). Of the six productivity measures the team developed, four were directly related to dollar units. The research team concluded the

primary purpose of performance measurements was for self-analysis. Kaneda and Wallett also came to the conclusion that the measurements were not comparable among other sections or squadrons (33:79-80). This is consistent with the finding of Baumgartel and Johnson and places them in Table 2 also under standard costing systems.

Productivity and Performance Measurements in the Air Force

Operational Performance Measurement.

The productivity of operational units is of major importance to the Air Force. Various productivity measurements or performance measurements are used.

Sortie rate, how many aircraft missions can be flown, is one frequently used operational measurement. Is this truly an important measurement? It does satisfy the basic definition of actual output over expected output but it does not completely satisfy the other relationships previously stated. Operational performance tends to be easier to measure because results usually are quantifiable, such as in sortie rates, alert rates, and flying hours.

Support Measurement.

Support areas have difficulty establishing performance measurement indicators because they may not have a direct, readily apparent result on how well the mission has been performed. Many support areas are customer service oriented or at least include customer service as part of their mission. These customer service organizations are best described by

Leon Greenberg, the former Executive Director of the National Commission of Productivity, Washington, D.C..

> There are thousands of establishments in many different industries producing a service . . .often with a recognizable physical form and yet extremely difficult to quantify and measure. The reason frequently given for nonmeasurement of these activities is that there is no definable, quantifiable unit of output. The more likely reason is that there is often more than one way to identify or define output of an establishment and it is difficult to make a choice. Another important factor is that each unit of output is often subject to a wide band of quality, and the quality is difficult to evaluate and measure. (27:40)

Air Force Manual 77-310, Volumes I,II,IV.

Air Force Manual 77-310 is the series concerning the operation and maintenance of Air Force vehicles and equipment. According to this manual:

> 37% of the total maintenance and operations costs expended to support the Air Force vehicle fleet or over $117 [million each year are labor costs]. Labor cost is used at base level to support vehicle repair decisions and budgets. (17:22)

This manual also provides that if required, the VOC percentage goals can be used to "support MAJCOM management programs" (17:12). AFM 77-310 permits major command management to tailor VOC goals by base, dependent on mission and operating environment (17:12). This recognizes the need for mission and environmental factors in the design of performance measurement standards.

USAF Performance Measurement Programs

One example of support performance measurement design is embodied by the recently implemented READY RESOURCE program.

This program focused on the response time of a mission support section to a maintenance action. This program repositioned over 60,000 parts to obtain a response time of under 15 minutes (40:1). To obtain data that allows such repositioning to occur, a database must be established. The database in this case is the Maintenance Data Collection System (MDC) (37:1). Data is input via a Computer Aided Maintenance System terminal (37:1). In order for this data to be useful, it must be accurate. The recent improvement was partly the result of constant reminding of personnel to input correct data (37:1).

Most middle managers in the Air Force are under such pressure to enhance their reported performance, the reasons for doing so are sometimes lost. The middle manager may wonder why the data is so important that his or her job may depend on it. Many times the middle level manager is force to "game" the system, to raise reported performance measurements to the standard or above. In these cases the measurement standard should be reevaluated based on its attainability.

Measuring Air Force performance has become increasingly important because of the economical constraints currently facing the Air Force. One of the largest areas of concern is contractor performance. The Air Force cannot afford to buy B1-B bombers that cannot fly, or contract for transportation services that are not performed well (6:38).

One of the latest attempts at measuring contractor performance is to include past performance of the contractor. An attempt is being made to eliminate the "blind government

concept" where the government will let contracts based on what the contractor promises. An example of this is the 1988 purchase of Tomahawk missiles. The Navy split this purchase between General Dynamics and McDonnell Douglas but reserved 10% of the buy for the company with the best quality of product. Unfortunately, the quality was relatively equal in both companies. Instead, the Navy used price as the selection factor and bought 70% from General Dynamics. This is an indication of future procurement changes and a most needed one (12:38-42).

The Air Force has made a recent proposal to develop a contract performance assessment reporting (CPAR) system. This database would cover all past performance of contractors working for the government. This database then would be used to assist in the awarding of future contracts (12:38-42).

General Alfred G. Hansen recently reorganized the Air Force Logistics Command's quality program. General Hansen feels performance measurements should relate to how long something works after it is fixed and not how many defects were reported on delivery (30:3). This would indicate that in transportation, the vehicle maintenance branch should measure how long a vehicle remains operational after repair instead of how many times the vehicle was turned in and a work order completed.

Future Trends

A new direction in operations management seems to indicate the amount of throughput time a product takes in the transformation process should be a unifying productivity measurement. This of course parallels the trends of just-in-time inventory management and other Japanese manufacturing philosophies. The reduction of throughput time seems to be a less expensive method than say increasing productivity by automation, considering the costs of automation. For example, General Motors has invested billions in the automation of its plants to improve productivity. On the other hand, the joint GM-Toyota Nova plant in Fremont, California has yielded equally impressive results by reduction of throughput time rather than a heavy investment in automation (39:717-718).

Sink notes this new direction in performance and productivity by stating,

> We are evolving from an area in which control was the major focus of most measurement and evaluation systems to an era in which development, commitment, involvement, etc . . . , will be the major focus. (41:86)

This statement agrees with other research reviewed in that performance measurement should be constructive and not punitive in nature (10:44).

Summary

The performance measure is a very important tool for today's managers in both civilian and military sectors. There are many conflicting theories on the establishment of

performance measures. There are no definitive models or frameworks in the field of performance or productivity. In the first section, special attention was directed towards those works advocating performance measurement criteria similar to A.T. Kearney's list of seven performance criteria.

In the second section, special attention was focused on those works examining the actual input/output, total productivity (capital+labor), standard costing, and inventory sources of performance measures. These included those works examining the benefits of a cost based measurement system, were transportation related, advocated the tracking of resources, and concerned nonprofit organizations. These sources are found listed in Table 2.

The third section contained an examination of current Air Force progress in operational and support areas of performance measurement. This section included a few new promising programs such as the READY RESOURCE maintenance program and the Contractor Performance Assessment Reporting System (CPARS). In the final section, future trends demonstrate the emergence of throughput time as a critical measure of operations management effectiveness.

Establishing effective performance measurement indicators which incorporate the seven characteristics from the A.T. Kearney study, and regularly evaluating these performance measurement indicators should be a goal in every organization. The benefits can far outweigh the costs. Chapter III examines the method of research used for this study.

III. Research Design and Methodology

Introduction

Given that the measurement of performance and productivity is a vital part of a good organization, the determination of whether Air Force transportation squadrons have effective performance measurement indicators will attempt to be made.

The central management question addressed was, "Do Air Force transportation squadrons have effective performance measurement indicators?" These are the research questions that guided the research approach.

1. What are the base-level transportation performance measurement indicators?
2. What are essential characteristics of performance measurement indicators?
3. What are the actual perceptions of transportation officers towards the base-level performance measurements using a reference standard of seven currently accepted criteria?

The intent of this chapter is to convey the research design executed during the course of this study. This research design furnishes a clear and replicable investigative effort that sufficiently answers the investigative questions stated in Chapter I. Literature relevant to the subject matter is reviewed and expert personnel interviewed to more accurately define the problem and provide historical material for the research effort. Next, the population for the study is chosen, described, and a plan for sampling this target is

formed. A survey instrument is designed and a plan constructed for collecting and analyzing the data.

Justification

A questionnaire format was chosen, as questionnaires are best used to measure interests, attitudes, beliefs, feelings, and perceptions according to Emory (22:158). Additionally, the size of the selected population and the information needed was large enough to make the use of a telephone survey impractical. The amount of time required to perform a telephone survey would exceed available resources. Research shows quality of information deteriorates as the telephone survey process lengthens (21:55).

Research Question One

Research question one, "What are the base-level transportation performance measurement indicators?" was answered by gathering primary data from personal interviews as well as secondary data from the review of Air Force literature. The personal interviews and literature review are provided in separate sections because of their size, and because they apply to more than one research question.

Research Question Two

Question two, "What are essential characteristics of performance measurement indicators?" was answered by gathering secondary data during literature review process. The

literature review is provided in a separate section because it applies to more than one research question.

Research Question Three

Question three, "What are the actual perceptions of transportation officers towards the base-level performance measurements using a reference standard of seven currently accepted criteria?" was answered by gathering primary data from a representative sample that would reflect the perceptions of performance measurements within that population. Additional insight was provided by the key personal interviews.

A.T. Kearney's criteria (34:43) were placed in questions within the survey. Each question on performance measurement contained one performance indicator and asked the respondents to evaluate the indicator based on the seven criteria.

Key Personnel Interviews

The Major Command Headquarters of each of the respective target major commands receive and maintain much of the performance measurement information. Because of this, the MAJCOMs of the target population was contacted. Recognized leaders in the field of Air Force transportation were contacted to answer investigative question one, "What are the base-level transportation performance measurement indicators? These indicators formed the basis of the questionnaire used to survey a representative sample of transportation officers.

Literature

An extensive review of the relevant literature was done concerning performance measurement, performance measurement systems, productivity measurement systems and physical distribution systems to answer investigative question two, "What are essential characteristics of performance measurement indicators." A comprehensive review of Air Force regulations, pamphlets, instructions, and other publications also was done to ascertain a response to question two. A Defense Technical Information Center (DTIC) literature search was accomplished to ensure the completeness of this literature review.

Target Population

The population selected was all transportation squadron commanders within the continental United States from the Strategic, Tactical, and Military Airlift Commands and selected transportation officers at the above mentioned major command headquarters, whose primary jobs focused on the management and maintenance of vehicle fleet resources. The questionnaire measured perceptions of performance measurement in both the vehicle operations and maintenance activities, and was tailored for application within the United States.

	Line	Staff
SAC	59	8
MAC	63	4
TAC	63	4

Figure 1. Population Matrix

Sampling Plan

Because of the relatively low number of officers within the population, a census of these individuals was attempted to get an accurate result. A total of 67 officers were surveyed.

Instrument.

The pilot questionnaire was sent to sample respondents. The test instrument was designed to answer investigative questions one and three. A seven point, from one to seven, Likert scale was used to rate each performance indicator based on the following seven criteria taken from the A.T. Kearney study on productivity discussed in Chapter II (34:43). These seven criteria were found by the A.T. Kearney consultants to be the most comprehensive list currently available (34:43).

To have an effective performance measurement indicator, the indicator should satisfy most of the seven general performance criteria in the Kearney study. These criteria are:

1. Validity
2. Coverage
3. Comparability
4. Completeness
5. Usefulness
6. Compatibility
7. Cost effectiveness (34:43)

This pilot questionnaire was tested on a small (5) convenience sample of squadron commanders. Based on their feedback, the questionnaire was modified and sent to the USAF Survey Control Office, Military Personnel Center, Randolph AFB, Texas. The center approved the questionnaire after

requesting minor changes. A survey control number and expiration date was issued. The requested changes and approval notice appears in Appendix B.

The questionnaire was divided into four major sections. Section I, consisting of questions one through twelve, requested background information and consisted primarily of demographic data.

Section II , consisting of questions thirteen through thirty-four, captured information about the perceptions of participants on the performance measurement indicators found in the Vehicle Integrated Management System (VIMS). The VIMS system was chosen as the primary set of maintenance performance measures by expert opinion obtained from key personnel interviews previously mentioned and research conducted in transportation regulations AFM 77-310, Volumes, I, II and IV, SACP 75-5, SACP 11-16, and 15AFP 75-1 (13)(14)(15)(16)(17)(18). In this section, respondents were asked to evaluate each indicator based on the seven criteria for effective performance measurement criteria, previously stated and obtained from the Kearney study (34:43). Brief definitions/explanations of each of the seven criteria were included on a cover sheet within the questionnaire booklet (see Appendix C). The respondent was asked to evaluate each indicator with respect to each of these criteria on a seven point Likert scale ranging from Strongly Agree to Strongly Disagree.

Section III, consisting of questions thirty-five through thirty-nine, provided information on five vehicle operations performance indicators using the same seven criteria on the same five point Likert scale. The particular performance measurement indicators were chosen by expert opinion obtained from the key personnel interviews previously mentioned, and research conducted in transportation regulations AFM 77-310, Volumes I, II, and IV, SACP 75-5, SACP 11-16, and 15AFP 75-1 (13)(14)(15)(16)(17)(18).

Section IV, consisting of questions forty through forty-five, provided perceptual information from respondents in five general areas. Based on the literature review, these five areas were commonly used in commercial performance measurement systems, but they were not addressed within the Air Force transportation measurement system. These five areas were manpower levels, environmental (geographic) conditions, costs, budget levels, and the use of performance factors to compare similar organizations. The survey ended with a request for narrative, open ended, responses from the participant concerning Air Force performance measurement indicators.

The overall management, specific research, and investigative questions were constructed using the research questions hierarchy (22:20). There are four levels in the hierarchy.

1. The management question-Do Air Force transportation squadrons have effective performance measurement indicators?

2. The research questions-What are the perceptions of Air Force transportation officers toward the effectiveness of transportation base-level performance measurement indicators in reflecting squadron performance?

3. The investigative questions-What are the base-level transportation performance measurement indicators? What are essential characteristics of performance measurement indicators? What are the actual perceptions of transportation officers towards the base-level performance measurements using a reference standard of seven currently accepted criteria?

4. The measurement questions-This consists of the questions found in the questionnaire. The primary data found here helps answer the previous questions.

Validation

The validity of the test instrument was established in two ways: a pilot test and expert opinion. An expert in the field of survey research examined the instrument. Then a content expert in the field of Air Force transportation reviewed the questionnaire. These experts found the questionnaire to be valid for measuring perceptions of Air Force officers concerning selected performance measurement indicators.

A pilot test of the instrument was accomplished by distributing six copies to current or recent squadron commanders by telefax machine. All of the respondents were transportation squadron commanders or recently had been in command of a transportation squadron. Four of the surveyed officers replied to the questionnaire, and minor changes were made to the test instrument based on the responses.

These two methods of survey validation via expert opinion and a pilot test are recommended by research experts (22:206).

These two approaches contributed to the construct validity of the subject questionnaire.

Data Collection Plan

The questionnaire was sent to the USAF Survey Control Office, Military Personnel Center, Randolph AFB, Texas for approval. The center approved the questionnaire after requesting minor changes, and a survey control number and expiration date was issued. The requested changes and approval notice appears in Appendix B.

The primary data collected from the survey consisted of perceptions. The perceptions are assumed to be valid if the meaning of the test instrument questions was clearly understood (i.e. the instrument was valid).

Pre-printed address labels were attached to full size envelopes to facilitate the distribution of the questionnaires. The envelope contained the questionnaire booklet and a preaddressed return envelope. The questionnaires were mailed as a block on June 27, 1989 from the Operations Office at building 641 of the Air Force Institute of Technology, Wright-Patterson AFB, Ohio to the target population.

The participation of individuals in the study was voluntary. In an effort to increase response rate, a postcard was mailed to the individuals ten days after the test instrument. A copy of the postcard appears in Appendix D. A second mailing of the survey instrument was completed seven

days after the postcard was sent. Personal phone calls were made to members of the sample population that had not returned the survey by five days after the second mailing. Anonymity was guaranteed to all participants to alleviate possible fears of retribution. A copy of the cover letter and questionnaire appear at Appendix C.

Data Analysis

The primary data recovered from this research was analyzed using Statistix: An Interactive Statistical Analysis Program for Microcomputers. The Statistix procedures used to analyze the data included: ONE, TWO, and MULTI-SAMPLE TESTS (Wilcoxen Rank Sum Test), and SUMMARY STATISTICS (Descriptive statistics, Frequency distributions, and Histograms) (42:6.12,9.4,9.5,9.7).

Statistical Design

The design of the test instrument used the Likert scale. The data obtained from this type of scale is ordinal in nature and is sometimes thought to be restricted to "descriptive statistics" such as percentages, means, modes and medians. There is support from some researchers that the use of ordinal data is just as accurate as that obtained from ratio or interval scales.

> In general, we are perfectly safe in calculating any statistic we want on any set of measurements that have the properties of an ordinal scale. There is definitive evidence that statistics calculated on ordinal measurements are just as reliable and meaningful as statistics calculated on interval or ratio scales of measurement. (3:309)

However, this research took the conservative approach to the ordinal-ratio data problem. Analysis relied on the use of nonparametric statistics because of the research assumptions of ordinal data.

The primary data obtained in this study was analyzed primarily using descriptive statistics means, modes, standard deviations and percentages. Demographics data and descriptive statistics were analyzed to find any possible correlations among the data. Relevant subgroups, of the population, were chosen for comparison purposes. Percentage responses from section II of the survey were analyzed and compared with the relative rankings from section IV of the survey as a means of validating the performance measurements perceived as most important by the respondents.

Frequency of response was used to determine the type of statistical test most appropriate for the data. The distribution of the responses indicated a nonparametric procedure should be used. The distribution of responses were skewed to the right as shown by the percentage response results from Table 8, the descriptive statistics response results in Table 9, the frequency response results in Table 12, and the histograms of the subgroups in Appendix B. A Wilcoxen Rank Sum Test was used because the assumption of normality could not be made and the use of a nonparametric procedure for determining a difference between the means of

two paired samples was indicated. Subgroups were tested to examine any differences in responses through use of the following procedure.

Wilcoxen Rank Sum Test

1. For Line officers and Staff officers responses.
2. For SAC officers and MAC officers responses.
3. For SAC officers and TAC officers responses.
4. For MAC officers and TAC officers responses.

H_o : The two populations are identical.

H_a : The two officer populations have different location parameters.

Test Statistic: $z = \frac{T-u_t}{\sigma}$

Where T is the sum of the ranks in sample 1.

Rejection Region: For a value of alpha = .05 two-tailed, z=1.645,

reject H if: $|z| > z(alpha)$

The Wilcoxen Rank Sum Test is a nonparametric alternative that requires fewer assumptions than a parametric test. When possible nonnormality of data is present in the form of skewness, the Wilcoxen Rank Sum Test is more informative than a comparable parametric alternative (38:187). Since both populations are 10 or larger, the use of a z statistic was allowed in the Wilcoxen Rank Sum Test (38: 183).

Using a 95 percent confidence level (alpha=.05), a comparison was made between the groups Line and Staff. Again using a 95 percent confidence level (alpha=.05), paired

comparisons of SAC, MAC, and TAC groups were accomplished. The results of these tests can be found in Tables 14 and 15. In the event of a large number of tied ranks a z score with continuity correction should be used (38:183). This score is provided in Tables 14 and 15.

Summary

In summary, the research questions were answered by the analysis of primary and secondary data. The primary data was gathered by key personal interviews and the use of a questionnaire on the target population. The secondary data was gathered from a review of military and commercial literature. The analysis was done using descriptive statistics, frequencies, percentages, and relative rankings. A statistical analysis also was done on chosen subgroups of the population to detect differences.

IV. Results

Introduction

There were 41 surveys (61%) returned by the established suspense date for analysis out of a population of 67. A complete listing of raw data is retained by AFIT/LSB. Although 55 total surveys were returned (82%), only 41 made the suspense date.

The remaining 14 surveys were excluded from data analysis because of time contraints and possible bias in the data because of late responses. These surveys were not returned before the suspense date and were not analyzed because of time restrictions. The possible presence of response bias was another reason not to include this data. Upon cursory examination four of the surveys contained the same response for almost every question in the survey.

An analysis was made of the nonrespondents and late respondents. The 26 officers comprising this group fell into three primary categories. More Captains did not respond than any other rank (10). More TAC personnel did not respond than any other command (11). This was determined by comparing the mailing list with the respondents. The percentages of nonrespondents in proportion to the number surveyed is provided in Table 3.

Additional information was available on the late respondents from the demographic data. These also fell into

Table 3

Nonresponse and Response Data

Category	Percentage Nonresponse	Percentage Response
Captains	38	27
Majors	34	37
Lieutenant Colonels	27	34
SAC	39	37
MAC	19	17
TAC	42	44
Line	85	83
Staff	15	17

three distinct groups. Most late respondents had less than two years in transportation, stated they had little or no experience in ground transportation, responded as having a primary AFSC other than transportation, or had some combination of the three.

Overview

The results of the data are grouped into six categories. They are listed below.

1. Population Demographics
2. Percentage Response Results
3. Descriptive Statistics of Chosen Subgroups
4. Frequencies of Average Responses by Chosen Subgroups
5. Relative Rankings of Performance Measurements
6. Wilcoxen Rank Sum Test of Chosen Subgroups

Note: 1. Complete copies of the final survey instrument are in Appendix C. 2. All survey questions are referred to as "Q" followed by the number and the letter of the question. 3. For comparison purposes, line and staff refer to the categories, line officer and staff officer. 4. SAC, MAC, and TAC refer to those respondents from the Strategic Air Command, Military Airlift Command, and the Tactical Air Command, respectively. 5. All percentage responses are "relative" based on actual responses in that category divided by the total number of responses. 6. All statistical output not listed in tables within the text can be found in Appendix E.

Population Demographics

Table 4 illustrates the demographic distribution of the respondents obtained from section I of the survey. The category rank consisted of three relatively equal components, Captain, Major, and Lieutenant Colonel, with Major being the largest with 37 percent. Because the majority of respondents were squadron commanders the expected rank of respondent was Major or Lieutenant Colonel.

The average number of years in a command or staff position at 3.6 was approximately one-half the average number of years in transportation at 7.6 years. Of those reporting the number of squadron personnel, the average number was 153.4. The largest squadron consisted of 320 personnel and the smallest had 71 personnel.

The officers surveyed consisted of 83 percent line officers and 17 percent staff officers. Nearly 84 percent of the officers had completed a graduate degree, and an additional 10 percent had completed graduate level work. Only 5.5 percent had no graduate work at all. The average age of these officers was 39. The respondents primarily came from the Tactical Air Command (TAC) with 44 percent followed by the Strategic Air Command (SAC) with 39 percent.

Percentage Response Results

The responses to section II and III represent the primary areas of interest in this study. The response data from section II are contained in Tables 5 through 11 and are

categorized according to the seven criteria from the A.T. Kearney study (34:43). The response data from section II is entitled "Evaluation of Vehicle Integrated Management System Performance Indicators." All question answers are arranged in a five point Likert scale with the following possible responses.

1. Strongly Agree
2. Agree
3. Neutral
4. Disagree
5. Strongly Disagree

Questions 13 through 39 were designed to ask respondents whether each performance measurement in the Vehicle Integrated Management System (VIMS) meets the seven criteria from the A.T. Kearney study summarized in Chapter II (34:43). The criteria are listed below for convenience.

a. Validity
b. Coverage
c. Comparability
d. Completeness
e. Usefulness
f. Compatibility
g. Cost Effectiveness (34:43)

Over 15.4 percent of the respondents answered with a response of 1, or Strongly Agree, to all of the questions in section II. A much higher response of 41 percent of the respondents answered with an response 2, or Agree, to all of the questions in section II. A considerably lower number of 24 percent answered 3, or Neutral, to those questions. Over 13.4 percent of the respondents gave a response of 4, or Disagree to these questions. The lowest percentage of the

Table 4

Demographic Response Distribution

RANK	0% LT	27% CAPT	37% MAJ	34% LTC
PRIMARY AFSC	73% 6016	22% 6054	5% OTHER	
AVE.NO.YRS. COMMAND OR STAFF POSITION	3.6			
AVE.NO.YRS. TRANS.	7.6			
AVE.NO.PERS. IN SQUADRON	153			
TYPE OFFICER	83% LINE	17% STAFF		
AVE. AGE	39			
LEVEL ED.	5.5% COMPLETED COLLEGE	10.7% SOME GRADUATE WORK	83.7% GRADUATE DEGREE	
MAJCOM	39% SAC	17% MAC	44% TAC	

population answered with responses of 5, or Strongly Disagree, with 6.5 percent.

Individual percentage responses are categorized by characteristic and question number in Tables 5 through 11. The highest percentage response of 1's tied at 39 percent for question 21 characteristics A, C, and F. The performance measure was "Percentage VDM" and the characteristics were Validity, Comparability, and Usefulness.

The highest percentage, 63 percent, of agree responses was found in questions, Q24-E, Q25-E, Q28-E. All questions concerned the characteristic, Usefulness. The performance measurements were "Percentage Hours VOC for three different

Table 5

Percent Response Distribution for All Respondents by Number of Question and Characteristic "VALIDITY"

Question	RESPONSE Performance Measurement	%1	%2	%3	%4	%5
13-a.	1. Ave. Cost Unit-Reg.	7	29	29	17	15
14-a.	2. Safety Insp. Overdue-Reg.	20	44	7	22	7
15-a.	3. Sched. Insp. Overdue-Reg.	22	41	12	17	7
16-a.	4. No. Work Orders Opened	17	34	17	24	7
17-a.	5. Hours Delay Code C	7	39	41	12	0
18-a.	6. Percent Direct Labor	17	39	10	15	20
19-a.	7. Percent Indirect Labor-Pr.	17	27	22	15	20
20-a.	8. Percent Indirect Labor-NPr.	12	32	20	15	22
21-a.	9. Percent Hours VDM-Reg.	37	54	7	2	0
22-a.	10. Percent Hours VDP-Reg.	34	49	10	7	0
23-a.	11. Percent Hours VOC-Reg.	34	49	10	7	0
24-a.	12. " " " -B&K	32	49	12	2	2
25-a.	13. " " " -D&M	24	51	17	5	2
26-a.	14. " " " -W	12	49	20	10	2
27-a.	15. " " " -E (463L)	32	44	15	7	2
28-a.	16. " " " -E (MHE)	27	46	17	7	2
29-a.	17. " " " -Firetr.	29	44	12	12	2
30-a.	18. " " " -Refuel.	32	41	15	10	2
31-a.	19. " " " -C&L	22	44	24	5	2
32-a.	20. " " " -Nonreg.	7	44	27	5	15
33-a.	21. VDM/Direct Labor Per.-Reg.	12	37	29	17	5
34-a.	22. Ave. Cost Admin. Fleet	7	32	29	20	7
35-a.	23. Taxi Ave. Resp. Time	12	41	27	20	5
36-a.	24. Veh. Op. Care-Wing	20	34	24	17	5
37-a.	25. Veh. Op. Care-Trans.	24	39	24	7	5
38-a.	26. No. Veh. Misuses	7	32	32	17	12
39-a.	27. No. Veh. Abuses	17	15	24	12	10

*Percentage may not add to 100 because of rounding.

categories of vehicles, B & K, D & M, and E (MHE). These vehicle codes represent refueling trucks, buses and truck tractors for B & K. Towing, servicing vehicles, base maintenance vehicles and material handling equipment are included in the D & M codes. The code E (MHE) includes material handling equipment.

The highest percentage of neutral responses was found in Q17-E, Q20-D, and Q35-D at 46 percent. The performance

Table 6

Percent Response Distribution for All Respondents by Number of Question and Characteristic "COVERAGE"

Question		RESPONSE Performance Measurement	%1	%2	%3	%4	%5
13-b.	1.	Ave. Cost Unit-Reg.	7	41	24	20	7
14-b.	2.	Safety Insp. Overdue-Reg.	22	39	27	10	2
15-b.	3.	Sched. Insp. Overdue-Reg.	17	44	24	10	5
16-b.	4.	No. Work Orders Opened	7	37	27	22	7
17-b.	5.	Hours Delay Code C	7	44	35	7	2
18-b.	6.	Percent Direct Labor	10	27	22	20	22
19-b.	7.	Percent Indirect Labor-Pr.	10	17	34	27	12
20-b.	8.	Percent Indirect Labor-NPr.	5	20	34	27	15
21-b.	9.	Percent Hours VDM-Reg.	24	49	17	10	0
22-b.	10.	Percent Hours VDP-Reg.	20	59	20	7	0
23-b.	11.	Percent Hours VOC-Reg.	24	49	17	10	0
24-b.	12.	-B&K	24	39	24	10	2
25-b.	13.	-D&M	12	49	29	7	2
26-b.	14.	-W	12	46	22	12	7
27-b.	15.	-E (463L)	20	51	20	4	5
28-b.	16.	-E (MHE)	17	54	20	7	2
29-b.	17.	-Firetr.	27	44	17	10	2
30-b.	18.	-Refuel.	24	46	20	7	2
31-b.	19.	-C&L	17	49	24	7	2
32-b.	20.	-Nonreg.	7	46	27	5	15
33-b.	21.	VDM/Direct Labor Per.-Reg.	10	27	32	27	5
34-b.	22.	Ave. Cost Admin. Fleet	7	29	39	22	2
35-b.	23.	Taxi Ave. Resp. Time	2	39	32	20	7
36-b.	24.	Veh. Op. Care-Wing	7	46	24	17	5
37-b.	25.	Veh. Op. Care-Trans.	20	37	29	10	5
38-b.	26.	No. Veh. Misuses	2	27	37	17	17
39-b.	27.	No. Veh. Abuses .	12	34	29	15	10

*Percentage may not add to 100 because of rounding.

measurements were "Hours Delay Code C", "Percent Indirect Labor-Nonproductive", and "Taxi-Average Response Time." The characteristics were Completeness for the first question and Usefulness for the second and third.

The highest percentage of disagree responses was found in Q13-C, "Average Cost per Unit", at 34 percent. The characteristic category was Comparability.

Table 7

Percent Response Distribution for All Respondents by Number of Question and Characteristic "COMPARABILITY"

Question		RESPONSE Performance Measurement	%1	%2	%3	%4	%5
13-c.	1.	Ave. Cost Unit-Reg.	5	34	22	34	5
14-c.	2.	Safety Insp. Overdue-Reg.	27	39	17	12	5
15-c.	3.	Sched. Insp. Overdue-Reg.	29	37	17	12	5
16-c.	4.	No. Work Orders Opened	7	39	24	24	5
17-c.	5.	Hours Delay Code C	7	35	37	12	5
18-c.	6.	Percent Direct Labor	15	41	10	20	15
19-c.	7.	Percent Indirect Labor-Pr.	15	27	27	20	12
20-c.	8.	Percent Indirect Labor-NPr.	10	24	29	22	15
21-c.	9.	Percent Hours VDM-Reg.	37	34	12	17	0
22-c.	10.	Percent Hours VDP-Reg.	32	34	15	20	0
23-c.	11.	Percent Hours VOC-Reg.	32	34	15	20	0
24-c.	12.	-B&K	27	39	17	15	2
25-c.	13.	-D&M	17	44	17	20	2
26-c.	14.	-W	15	39	17	24	5
27-c.	15.	-E (463L)	22	37	12	22	7
28-c.	16.	-E (MHE)	20	39	12	24	5
29-c.	17.	-Firetr.	29	29	20	15	7
30-c.	18.	-Refuel.	27	34	20	17	2
31-c.	19.	-C&L	20	37	24	15	5
32-c.	20.	-Nonreg.	10	29	37	12	12
33-c.	21.	VDM/Direct Labor Per.-Reg.	12	34	24	24	5
34-c.	22.	Ave. Cost Admin. Fleet	10	37	20	27	7
35-c.	23.	Taxi Ave. Resp. Time	2	44	20	17	17
36-c.	24.	Veh. Op. Care-Wing	12	39	24	20	5
37-c.	25.	Veh. Op. Care-Trans.	20	29	27	22	2
38-c.	26.	No. Veh. Misuses	2	29	27	27	15
39-c.	27.	No. Veh. Abuses	15	32	15	24	15

*Percentage may not add to 100 because of rounding.

The highest percentage of strongly disagree responses was in Q18-B and Q20-A. The performance measurements were "Percent Direct Labor" and "Percent Indirect Labor-nonproductive." The characteristics were Validity and Coverage, respectively.

Section IV of the survey had the purpose of obtaining the perceptions of the population on the desirability of including manpower, budget, environment, and actual cost figures in their performance measurement indicators. Section IV surveyed

Table 8

Percent Response Distribution for All Respondents by Number of Question and Characteristic "COMPLETENESS"

Question		RESPONSE Performance Measurement	%1	%2	%3	%4	%5
13-d.	1.	Ave. Cost Unit-Reg.	10	41	29	15	5
14-d.	2.	Safety Insp. Overdue-Reg.	10	46	32	10	2
15-d.	3.	Sched. Insp. Overdue-Reg.	15	41	32	10	2
16-d.	4.	No. Work Orders Opened	7	41	27	17	7
17-d.	5.	Hours Delay Code C	5	41	39	12	2
18-d.	6.	Percent Direct Labor	15	29	22	17	17
19-d.	7.	Percent Indirect Labor-Pr.	12	22	32	20	15
20-d.	8.	Percent Indirect Labor-NPr.	5	15	46	22	12
21-d.	9.	Percent Hours VDM-Reg.	27	49	17	5	2
22-d.	10.	Percent Hours VDP-Reg.	24	44	20	10	2
23-d	11.	Percent Hours VOC-Reg.	20	51	20	10	2
24-d.	12.	-B&K	20	51	20	10	2
25-d.	13.	-D&M	12	56	24	2	5
26-d.	14.	-W	10	56	20	7	7
27-d.	15.	-E (463L)	24	49	17	5	5
28-d	16.	-E (MHE)	20	51	17	7	5
29-d.	17.	-Firetr.	24	49	20	2	5
30-d.	18.	-Refuel.	24	49	20	2	5
31-d.	19.	-C&L	17	46	24	7	5
32-d.	20.	-Nonreg.	5	49	29	2	15
33-d.	21.	VDM/Direct Labor Per.-Reg.	10	34	29	15	12
34-d.	22.	Ave. Cost Admin. Fleet	7	37	24	22	10
35-d.	23.	Taxi Ave. Resp. Time	2	39	46	10	2
36-d.	24.	Veh. Op. Care-Wing	10	37	37	15	2
37-d.	25.	Veh. Op. Care-Trans.	17	37	34	10	2
38-d.	26.	No. Veh. Misuses	5	24	32	27	12
39-d.	27.	No. Veh. Abuses	15	27	34	17	7

*Percentage may not add up to 100 because of rounding.

respondents on the comparability of base-level transportation organizations using the performance management indicators in Q13-39.

The results of Section IV were the most dramatic of the survey. In Q41, the respondents responded 83 percent in favor and 12 percent against manpower being used. Over 70 percent responded in favor of environment being considered in performance measurement design, in Q42. In Q43, the respondents responded 76 percent in favor and 10 percent

Table 9

Percent Response Distribution for All Respondents by Number of Question and Characteristic "USEFULNESS"

Question		RESPONSE Performance Measurement	%1	%2	%3	%4	%5
13-e.	1.	Ave. Cost Unit-Reg.	10	27	37	15	10
14-e.	2.	Safety Insp. Overdue-Reg.	32	49	10	10	0
15-e.	3.	Sched. Insp. Overdue-Reg.	34	46	12	7	0
16-e.	4.	No. Work Orders Opened	10	41	22	20	10
17-e.	5.	Hours Delay Code C	7	39	46	7	0
18-e.	6.	Percent Direct Labor	10	46	20	15	10
19-e.	7.	Percent Indirect Labor-Pr.	12	34	27	20	7
20-e.	8.	Percent Indirect Labor-NPr.	23	29	22	22	20
21-e.	9.	Percent Hours VDM-Reg.	37	54	7	2	0
22-e.	10.	Percent Hours VDP-Reg.	29	56	10	5	0
23-e.	11.	Percent Hours VOC-Reg.	27	34	29	5	2
24-e.	12.	-B&K	22	63	10	5	0
25-e.	13.	-D&M	12	63	17	5	2
26-e.	14.	-W	10	54	27	7	2
27-e.	15.	-E (463L)	20	61	20	0	0
28-e.	16.	-E (MHE)	17	63	17	2	0
29-e.	17.	-Firetr.	22	59	12	7	0
30-e.	18.	-Refuel.	24	59	12	5	0
31-e.	19.	-C&L	15	59	20	7	0
32-e.	20.	-Nonreg.	7	41	27	10	15
33-e.	21.	VDM/Direct Labor Per.-Reg.	15	39	20	17	10
34 e.	22.	Ave. Cost Admin. Fleet	12	32	24	22	10
35 e.	23.	Taxi Ave. Resp. Time	7	51	27	2	12
36 e.	24.	Veh. Op. Care-Wing	17	41	29	7	5
37-e.	25.	Veh. Op. Care-Trans.	27	34	29	5	2
38-e.	26.	No. Veh. Misuses	10	29	32	15	15
39-e.	27.	No. Veh. Abuses	24	29	27	12	7

*Percentages may not add to 100 because of rounding.

against the presence of actual cost figures in dollars in performance measurement indicators. A majority of 61 percent in favor and 17.4 percent responded against, in Q44, the use of squadron budget levels in performance measurement design. Question 45 yielded decisive results with a 49 percent against and 22 percent for the use of the performance measurement in Q13-39 being used to compare squadron performance across all bases.

Table 10

Percent Response Distribution for All Respondents by Number of Question amd Characteristic "COMPATIBILITY"

Question		RESPONSE Performance Measurement	%1	%2	%3	%4	%5
13-f.	1.	Ave. Cost Unit-Reg.	5	49	34	7	5
14-f.	2.	Safety Insp. Overdue-Reg.	22	51	17	7	2
15-f.	3.	Sched. Insp. Overdue-Reg.	20	51	20	7	2
16-f.	4.	No. Work Orders Opened	10	56	22	7	5
17-f.	5.	Hours Delay Code C	7	54	37	0	2
18-f.	6.	Percent Direct Labor	15	51	15	15	5
19-f.	7.	Percent Indirect Labor-Pr.	12	41	24	17	5
20-f.	8.	Percent Indirect Labor-NPr.	7	39	29	15	10
21-f.	9.	Percent Hours VDM-Reg.	32	51	15	2	0
22-f.	10.	Percent Hours VDP-Reg.	27	56	15	2	0
23-f.	11.	Percent Hours VOC-Reg.	27	56	15	2	0
24-f.	12.	-B&K	24	51	17	5	2
25-f.	13.	-D&M	15	61	17	5	2
26-f.	14.	-W	15	59	20	5	2
27-f.	15.	-E (463L)	24	54	17	2	2
28-f.	16.	-E (MHE)	22	56	17	2	2
29-f.	17.	-Firetr.	29	49	15	5	2
30-f.	18.	-Refuel.	27	51	15	5	2
31-f.	19.	-C&L	20	54	20	5	2
32-f.	20.	-Nonreg.	7	46	29	2	15
33-f.	21.	VDM/Direct Labor Per.-Reg.	15	37	27	17	5
34-f.	22.	Ave. Cost Admin. Fleet	7	49	24	17	2
35-f.	23.	Taxi Ave. Resp. Time	12	37	34	12	5
36-f.	24.	Veh. Op. Care-Wing	10	37	27	22	5
37-f.	25.	Veh. Op. Care-Trans.	15	39	27	15	5
38-f.	26.	No. Veh. Misuses.	5	27	37	17	15
39-f.	27.	No. Veh. Abuses .	10	39	32	10	10

*Percentages may not total to 100 because of rounding.

Descriptive Statistics of Chosen Subgroups

Table 12 gives decsriptive statistics of some important subgroups chosen from the demographic data. As stated in Chapter III, comparison of line and staff officers was of importance in this study. The comparisons of major commands, SAC, MAC, TAC, were of equal importance.

The highest mean response between line and staff officers was staff officers with 2.7. The staff officers range of

Table 11

Percent Response Distribution for All Respondents by Number of Question and Characteristic "COST EFFECTIVENESS"

Question		RESPONSE Performance Measurement	%1	%2	%3	%4	%5
13-g.	1.	Ave. Cost Unit-Reg.	7	29	37	20	7
14-g.	2.	Safety Insp. Overdue-Reg.	12	54	22	7	5
15-g.	3.	Sched. Insp. Overdue-Reg.	20	49	22	7	2
16-g.	4.	No. Work Orders Opened	2	41	24	17	15
17-g.	5.	Hours Delay Code C	7	34	44	12	2
18-g.	6.	Percent Direct Labor	10	32	24	15	20
19-g.	7.	Percent Indirect Labor-Pr.	10	20	37	20	15
20-g.	8.	Percent Indirect Labor-NPr.	7	17	32	20	24
21-g.	9.	Percent Hours VDM-Reg.	22	54	17	2	5
22-g.	10.	Percent Hours VDP-Reg.	20	49	22	2	5
23-g.	11.	Percent Hours VOC-Reg.	22	54	17	2	5
24-g.	12.	-B&K	20	51	20	7	2
25-g.	13.	-D&M	17	49	20	10	5
26-g.	14.	-W	12	51	17	15	5
27-g.	15.	-E (463L)	20	56	12	7	5
28-g.	16.	-E (MHE)	17	59	12	7	5
29-g.	17.	-Firetr.	24	49	15	7	5
30-g.	18.	-Refuel.	22	54	12	7	5
31-g.	19.	-C&L	17	51	20	10	2
32-g.	20.	-Nonreg.	5	41	22	10	22
33-g.	21.	VDM/Direct Labor Per.-Reg.	10	34	24	20	12
34-g.	22.	Ave. Cost Admin. Fleet	5	34	39	15	7
35-g.	23.	Taxi Ave. Resp. Time	5	37	41	7	10
36-g.	24.	Veh. Op. Care-Wing	12	22	41	17	7
37-g.	25.	Veh. Op. Care-Trans.	12	34	37	12	5
38-g.	26.	No. Veh. Misuses.	5	24	44	12	15
39-g.	27.	No. Veh. Abuses	20	27	27	17	7

*Percentage may not add to 100 because of rounding.

response was also greater, from an average minimum of 1.7 to an average maximum of 4.0.

Between major commands, MAC responded with the highest mean response at 2.697 followed by TAC with 2.552 and SAC with 2.341. The MAC officers also responded with the widest range of response from an average minimum of 1.4 to an average maximum of 3.7.

Table 12

Descriptive Statistics of Chosen Subgroups

Subgroup	Mean	S.D.	N	Median	Minimum	Maximum
Line	2.444	4.130E-01	224	2.40	1.70	3.50
Staff	2.737	4.989E-01	224	2.70	1.70	4.00
SAC	2.341	4.308E-01	224	2.30	1.60	3.30
MAC	2.697	4.878E-01	224	2.70	1.40	3.70
TAC	2.552	4.221E-01	224	2.40	1.80	3.70

Frequencies of Responses of Chosen Subgroups

Table 12 gives frequencies of average responses for the subgroups line, staff, SAC, MAC, and TAC. Line officers have the highest number of average responses between .5 and 1.5 with 19. Staff officers follow with 5. Line officers also have the highest average number of responses between 1.5 and 2.5 with 177. Staff officers have more average responses between 2.5 and 3.5 with 78. No average responses were found in the range 4.5 to 5.5.

Relative Rankings of Performance Measurement Indicators

In section IV of the survey respondents were asked to rank the performance measurements within the survey (Q13-Q39) in order of importance from most to least for one through ten. There was some confusion over whether the wording meant the five best and the five worst or the top ten (which is what was meant). The purpose was to determine which performance measurements were felt to be the most important overall, and compare the average rank of those performance measurements with the average total score of each of those measurements from Q13-Q39.

Table 13

Frequency Distribution of Chosen Subgroups

Line

Value	N	
1	19	|****
2	177	|**********
3	28	|******

Total 224

Staff

Value	N	
1	5	|*
2	140	|**********
3	78	|**********
4	1	|

Total 224

SAC

Value	N	
1	51	|**********
2	151	|**********
3	22	|******

Total 224

MAC

Value	N	
1	7	|**
2	144	|**********
3	73	|**********

Total 224

TAC

Value	N	
1	11	|**
2	177	|**********
3	36	|*******

Total 224

Table 14

Cumulative Number of Times Performance Measurement Occurs Within Each Ranking

Performance Measurement	Rank									
	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
Ave. Cost Unit-Reg	1	1	1	2	3	3	4	5	6	7
Safety Insp. O/D-Reg.	4	5	7	10	12	13	15	15	15	15
Sched. Insp. O/D-Reg.	0	4	5	6	9	11	13	14	14	14
No. Wo. Orders Opened	1	1	2	3	3	6	7	9	10	10
Hours Delay Code C	0	0	0	1	2	2	2	4	5	7
Percent Direct Labor	0	1	4	7	8	8	9	11	13	13
Percent In. Labor-Pr..	0	0	0	0	1	1	1	3	3	4
Percent In. Labor-NPr.	0	0	0	2	2	3	5	6	7	7
Percent Hours VDM-Reg.	7	13	16	17	17	18	18	19	21	21
Percent Hours VDP-Reg.	2	8	12	14	15	15	16	17	17	20
Percent Hours VOC-Reg.	9	9	13	13	14	15	15	15	15	16
" " -B&K	0	0	0	1	1	3	3	3	6	8
" " -D&M	0	0	0	0	2	2	2	4	4	4
" " -W	0	0	0	0	1	1	1	1	2	2
" " -E (463L)	1	4	5	5	8	9	10	10	11	11
" " -E (MHE)	0	0	0	2	2	6	8	9	9	10
" " -Firetr.	0	2	2	3	8	10	13	14	16	16
" " -Refuel.	0	0	0	1	2	5	8	13	14	15
" " -C&L	0	0	0	0	0	0	0	0	1	2
" " -Nonreg.	0	0	0	0	0	0	2	2	2	2
VDM/Direct Labor -Reg.	0	1	2	2	2	2	2	2	2	4
Ave. Cost Admin Fleet	0	0	2	2	2	4	4	6	7	9
Taxi Ave. Resp. Time	1	2	4	6	8	8	8	10	12	14
Veh. Op. Care-Wing .	2	2	6	9	11	11	12	13	15	16
Veh. Op. Care-Trans.	0	3	4	4	6	10	10	11	14	15
No. Veh. Misuses	1	1	1	2	3	5	7	10	10	10
No. Veh. Abuses	0	2	3	3	6	8	9	11	12	16

Unfortunately, almost 20 percent of the respondents failed to respond to this section, or responded but questioned the wording of the survey with respect to the rankings. The relative ranking of the measurements was recorded and placed in Table 14. The relative ranking should not change even if the question was misinterpreted. A comparison was made between the highest ranked performance measurements from the relative ranking and the highest average response of those measures from section II. This can be found in Chapter V.

Table 15

WILCOXEN RANK SUM TEST FOR CHOSEN SUBGROUPS

Test		Rank Sum	Sample Size	Average Rank	Normal Approximation	Two Tailed P value
1.	LINE	4.18E+04	224	186.7	41.653	0.000
	STAFF	5.87E+04	224	262.3		
	Calculated Z value=-6.17					
2.	SAC	4.00E+04	224	178.8	40.353	0.000
	MAC	6.05E+04	224	270.2		
	Calculated Z value=-7.48					
3.	SAC	4.36E+04	224	194.8	42.973	0.000
	TAC	5.69E+04	224	254.2		
	Calculated Z value=-4.86					
4.	MAC	5.48E+04	224	244.7	44.527	0.000
	TAC	4.576+04	224	204.3		
	Calculated Z value=3.3					

Test of the Chosen Subgroups by Wilcoxen Rank Sum Test

The results of the statistical tests performed on the chosen subgroups are provided in Table 15. As previously stated in Chapter III, the distribution of the response indicated a nonparametric procedure should be used. The distribution of responses were skewed to the right as shown by the percentage response results from Tables 4-11, the descriptive statistics response results in Table 12, the frequency response results in Table 12, and the histograms of the subgroups in Appendix E.

The Wilcoxen Rank Sum Test is a nonparametric alternative that requires fewer assumptions than a parametric test as stated in Chapter III. Using a 95 percent confidence level

(alpha=.05), a comparison was made between the groups Line, and Staff. The sum of the ranks was 40,183 for Line and 50,875 for Staff. The computed z value is -6.17.

Again using a 95 percent confidence level (alpha=.05), comparison between SAC and MAC yielded a sum of the ranks of 40,040 and 60,530 respectively. The resulting two tailed p value for normal approximation was zero. The normal approximation with continuity correction was 44.527. The computed z value was -7.48.

Using the same parameters, SAC and TAC yielded a sum of the ranks of 43,630 and 56,940. The resulting two tailed p value for normal approximation was zero. The normal approximation with continuity correction was 42.973. The computed z value was -4.86.

The same parameters for a comparison of MAC versus TAC yielded a sum of the ranks of 54,810 and 45,760 respectively. The resulting two tailed p value for normal approximation remained zero. The normal approximation with continuity correction was 44.527. The computed z value was 3.3.

Conclusion

The large amount of data collected resulted in a wide variety of results. The results indicated the respondents agreed in general that the performance measurements met the criteria necessary for good performance measurements. However, a considerable portion of the population perceived the performance measurements as not meeting the criteria. The

chosen subgroups were found to represent distinct divisions within the population.

Results were consolidated into demographics, percentage average response, descriptive statistics, frequencies, relative rankings, and the comparison of chosen subgroups with the Wilcoxen Rank Sum Test statistical test. Tables 3 through 15 were used to present the results of the data in a consolidated, more meaningful manner. An analysis of the data will be presented in Chapter V.

V. Analysis

Overview

There are many problems in measuring performance in a base-level transportation squadron that complicate productivity assessment of the unit. A transportation squadron is a nonprofit production and service organization. Performance indicators impact every facet of transportation and should be periodically evaluated.

This research determined the perception of respondents was that current transportation performance measurements, in general, meet the seven criteria for performance measurements from the A.T. Kearney study (34:43). However, there existed enough negative and neutral responses to warrant further analysis.

The results of this study are analyzed in the following sections. The results have been grouped into the following sections, the population demographics, the relative importance of the measures, comparisons of the chosen subgroups, and an analysis of the perceptions of the respondents from questions (Q) 41-45, Those questions cover the use of manpower, budget, environment, and actual cost in performance measurement design.

Population Demographics

The demographic data was used to ascertain general relationships between the demographic categories and the responses of the population. Each category found to be

significant is examined below. The summary of sample demographic characteristics shown in Table 3 consolidates the respondent data into a more easily analyzed form.

The category of "Primary AFSC" (Air Force Speciality Code) was designed to provide a subgroup of the population of officers who had made transportation a secondary AFSC. This would include rated officers as well as those from other logistics areas. No subgroup of analyzable size could be extracted from the data. An overwhelming majority of 95 percent of the respondents answered 6016 (field grade transportation position) or 6054 (company grade transportation officer) to this question. Only two officers (5%) responded as having other than a transportation primary AFSC. This subgroup was too small for significant analysis.

The category "Rank" showed a distribution similar to that of AFSC. The 6016 AFSC (transportation command position) which represented 73 percent of the population corresponded to the sum of the percents of Major and Lieutenant Colonel (71%). This sum is close to what would be expected because the population surveyed consisted of 76 percent line officers which were commanders and normally are majors or lieutenant colonels.

The category "Average Number of Years in Transportation" yield of 7.6 years did not have the same distribution as rank and primary AFSC. Since almost 75% of the population were of field grade rank, and practically all had named transportation as their primary AFSC, the expected number of years should be

roughly equivalent to the number of years in service. Upon further analysis of the demographic data, the average number of years in transportation for Majors and Lt. Colonels was 8.3. This clearly does not correspond to the number of years in service. A further breakdown by rank revealed the average years in transportation for Captain, Major, and Lt. Colonel was 10.75, 9.5, and 7.3. This indicates an inverse relationship of decreasing experience in transportation as rank increases in the respondents. This would also indicate many respondents had not been in transportation their entire careers. This can be explained by an influx of operational officers with no experience in transportation at the field grade level that reduces the mean.

The category "Type Officer," was designed to provide two distinct subgroups within the population for comparison purposes. The responses line or staff provided the two subgroups. The response also correlated with the category "Position." All line officers were also commanders.

Relative Importance of Performance Measurements

Criterion Responses.

The primary areas of interest of the study concerned the responses to questions in section II and III of the survey. The percentage of individual responses by number are provided in Tables 5-11. The distribution of the responses to the questions in sections II and III, (15.4%-1, 41%-2, 24%-3, 13.4%-4, 6.5%-5) indicates the respondents feel, in general,

the performance measurements in Q13-39 satisfy the general seven criteria from the A.T. Kearney study. There exists, however, a large neutral and disagree faction (42.9%) that warranting further examination. This may be an indication of disinterest or dissatisfaction with the present performance indicators.

The Strongly Agree and Agree percentages for each measurement were added together and averaged over all characteristics. The measurements having the fifteen highest averages are provided in Table 16. Then the five highest average agree responses for each characteristic are listed in Table 17. The Strongly Disagree and Disagree percentages for the three highest disagree responses were added together and averaged over all characteristics.

Vehicle Deadlined for Maintenance (VDM) averaged 80.1 percent response rate among agree and strongly agree over all seven of the criteria. This analysis of the data indicates the measurement is perceived to be the best performance measurement by the respondents. The measure also is ranked first on the relative ranking from Q40. Vehicle Deadlined for Parts (VDM) was second with 76.1. This corresponds exactly with the relative ranking of the top five measurements of Q40. The third highest overall was Percent VOC-E (463L). This measurement was not ranked in the top five measurements from Q40. The occurrence of a performance measurement not ranked on the relative scale of importance from Q40 was common.

Table 16

Average Total Agree Response Over Seven Characteristics

Performance Measurement	Ave. % Agree	Relative ranking
Percent VDM-Reg.	80.1	1
Percent VDP-Reg.	76.1	2
Percent VOC-E (463)	73.4	NR
Percent VOC-Reg.	73.3	3
Percent VOC-B&K	73.1	NR
Percent VOC-E (MHE)	72.6	NR
Percent VOC-Firetruck	72.4	6 (tie)
Percent VOC-Refueling	72	NR
Percent VOC-D&M	67.3	NR
Percent VOC-C&L	67.1	6 (tie)
Safety Insp. Overdue-Reg.	66.7	4
Sched. Insp. Overdue-Reg.	66.6	5
Percent VOC-W	62.9	NR
Veh. Op. Care Rating-Trans.	55.6	10
Percent Direct Labor	51	6 (tie)

*NR-Not ranked in the top five relative ranks from Q40.

However, the measurements were almost exclusively Vehicle Out of Commission (VOC) rates on different vehicle types. The general category VOC may be perceived as meeting the performance measurements characteristics but in the relative ranking from one to ten, the individual importance of the measurements is less. This is evidenced by the fact that if all VOCs for specific vehicle types were removed from Table 16, the top five performance measurements are exactly the same and in the same order as found in the relative rankings provided in Table 18. All performance measurements in the five highest ranked cumulative responses showed percentage results similar to those above.

The performance measurement which had the highest total disagree response (Disagree+Strongly Disagree) was Percent indirect labor-nonproductive in the characteristic "Cost

Table 17

Highest Total Agree Response for Each Characteristic

Performance Measurement	Ave. % Agree	Relative ranking
Validity		
1. Percent VOC-Reg.	91	3
2. Percent VDP-Reg	83	2
3. Percent VOC-B&K	81	NR
4. Percent VOC-E (463L)	76	6
5. Percent VOC-D&M	75	NR
Coverage		
1. Percent VDP-Reg.	79	2
2. Percent VDM-Reg.	73	1
Percent VOC-Reg.	73	3
4. Percent VOC-Firetruck	71	6
Percent VOC-E (463L)	71	6
Comparability		
1. Percent VDM-Reg.	71	1
2. Percent VOC-B&K	66	NR
Safety Insp. Overdue-Reg.	66	4
Sched. Insp. Overdue-Reg.	66	5
Percent VDP-Reg.	66	2
Completeness		
1. Percent VDM-Reg.	76	1
2. Percent VOC-E (463L)	73	6
Percent VOC-Firetruck	73	6
Percent VOC-Refueling	73	NR
5. Percent VDP,Percent VOC-B&K	71	2,NR,3
Percent VOC-Reg.		
Usefulness		
1. Percent VDM-Reg.	91	1
2. Percent VOC-B&K	85	NR
Percent VDP-Reg.	85	2
4. Percent VOC-Refueling	83	NR
5. Safety Insp. Overdue-Reg.	81	2,6,6
Percent VOC-E (463L)		
Percent VOC-Firetruck		
Compatibility		
1. Percent VDP-Reg.	83	2
Percent VOC-Reg.	83	3
Percent VDM-Reg.	83	1
4. Percent VOC-E (M	78	NR,6,6,NR,
Percent VOC-Firetruck	78	
Percent VOC-E (463L)		
Percent VOC-Refueling		
Cost Effectiveness		
1. Percent VOC-E (MHE)	76	NR
Percent VOC-E (463L)	76	6
Percent VOC-Refueling	76	NR
Percent VOC-Reg.	76	3
Percent VDM-Reg.	76	1

*NR=Not Ranked in the top five relative ranks from Q40.

Table 18

THE TOP TEN PERFORMANCE MEASUREMENTS IN THE TOP FIVE RANKS

RANK	MEASUREMENT	NUMBER OF OCCURRENCES
1.	9.	17
2.	10.	15
3.	11.	14
4.	2.	12
5.	3.	9
6.,7.,8.,9.	6.,15.,17.,23.	8
10.,11.	25.,27.	6

1. Average cost per unit-Registered
2. Safety inspection overdue-Registered
3. Scheduled inspection overdue-Registered
4. Number of work orders opened
5. Hours delay code C
6. Percent direct labor
7. Percent indirect labor-productive
8. Percent indirect labor-nonproductive
9. Percent hours VDM-Registered
10. Percent hours VDP-Registerd
11. Percent hours VOC-Registered
12. Percent hours VOC-B & K
13. Percent hours VOC-D & M
14. Percent hours VOC-W
15. Percent hours VOC-E (463L)
16. Percent hours VOC-E (MHE)
17. Percent hours VOC-Fire truck
18. Percent hours VOC-Refueling
19. Percent hours VOC-C & L
20. Percent hours VOC-Nonregistered
21. VDM/Direct labor percent -Registered
22. Average cost per mile-Admin fleet
23. Taxi-average response time
24. Vehicle operators care rating-Wing
25. Vehicle operators care rating-Transportation
26. Number of vehicle misuses
27. Number of vehicle abuses

Effectiveness." This measurement averaged 39 percent on the disagree side of neutral. This measurement is not present in top five cumulative rankings, lending further evidence to the credibility of the two types of evaluations. When compared with the 83 percent in favor of manpower being included in

performance measurement design, from Q41 in section IV of the survey, and a consensus among some researchers that labor be figured into performance measurements, this is a disturbing finding. Possible explanations could be respondents do not understand the construction or value of these particular labor measurements. Another explanation may be respondents do not perceive them as being important because they have little control over these figures. Finally, the respondents may feel that these performance measurements should be a part of measurement design but since they are not briefed regularly at the base level, or used to compare bases at the major command level, they lack importance.

The second highest total average disagree response over all seven criteria belongs to two measures of labor. Percent direct labor and Percent indirect labor-nonproductive measure 34 and 32 percent respectively, on the disagree side of neutral. The fact over one-third of respondents felt the three primary transportation labor measures do not meet the characteristics necessary for good performance measurements and also that the majority do not include the measures in their top ten most important measures indicates a serious problem with these measures. Whether the reasons are the ones stated above or there are other explanations does not lessen the seriousness of the problem.

Relative Importance by Ranking

The DOD productivity improvement program defines productivity, as a combination of efficiency and effectiveness. Efficiency meant to accomplish the mission correctly at least cost. Effectiveness meant to do those things at the right time. As specifically stated in DOD guidelines, "The efficiency with which organizations utilize all types of fund resources to accomplish their mission represents total resource productivity" (20:1). The importance of these definitions can be seen in the results of the relative ranking of performance measurements according to importance (Q40).

The results of Q40, the ranking of the performance measurement Q13-39 of the survey, are given in Table 14 and 18. The measurements Percent hours VDM-Registered, Percent hours VDP-Registered, and Percent hours VOC-Registered are respectively ranked first, second, and third. These measurements are the most general performance measurements of those listed that measure the inverse availability of the vehicle to the using organization. These measurements also measure the effectiveness of the organization in accomplishing the mission as the policy of the DOD states in the definition of productivity, a combination of efficiency and effectiveness (20:1). This would point to the transportation officer being able to recognize the measures that gauge customer service and operational support (here seen as Vehicle Out of Commission rate) as being more important to the mission than efficiency.

Since these numbers also are briefed frequently at staff meetings and used by major command to compare bases these items may also be perceived as important.

The fourth and fifth positions of highest cumulative response are Safety Inspections Overdue-Registered and Scheduled Inspections Overdue. Since these areas both refer to preventive maintenance, the customer service aspect of the measurement becomes obscured. The customer knows the vehicle is out of service during the time of inspection and may not be sympathetic to the reason. The customer may not consciously realize that the vehicle being out of service now, may keep it or his personnel in service, a higher percentage of the time in the future. This measurement could perhaps be modified to reflect the benefit to the customer, such as safety items repaired or worn items repaired.

Direct Labor is ranked the sixth highest cumulative measure. This is disturbing because of the emphasis placed on the use of labor as a measurement by researchers. This may reflect the lack of control the manager has over this measurement. Direct Labor is directly related to the amount of work in a shop. If the shop is operating at capacity direct labor will fall. Since the efficiency of the operation would in part be measured by this figure, the practice contradicts the DOD definition of productivity, a combination of effectiveness and efficiency (20:1). Other measures of efficiency also do not appear in the top ten highest cumulative ranks. These are Average Cost per Unit-Registered,

VDM/Direct Labor Percent-Registered, and Average Cost per Mile.

Taxi-average response time is also ranked the sixth highest cumulative measure. Since this measure reflects the level of customer service this is a measure of effectiveness. This measure is also a frequently briefed number at staff meetings and sometimes used to compare bases. This may indicate the measure is perceived as important because of these reasons as well as for measuring organization effectiveness.

Vehicle Operators Care Rating-Transportation is a very often briefed measure used to compare organizations at the same base and with other transportation squadrons. The measures are probably perceived as being important because of the reason stated above. The measure does not seem to reflect either the effectiveness or efficiency of the organization because it measurements are arbitrary and based on judgment alone. A high Vehicle Operator Care rating contributes little direct support to the mission (effectiveness) and the competition encourages the inefficient use of resources (cleaning supplies, wax, water, manpower). Anthony describes one of the problems of a nonprofit organization as an "unclear connection between benefits and costs" (1:42-44). Such is the case with this measure.

The other measurements appearing in the top ten rankings are a mixture of the VOC rates of specific vehicle types.

These are probably present for the same reasons stated above for VOC-Registered.

The measurements ranked as most important all concerned the effectiveness (ability to perform the mission) or operational aspect of transportation support. Efficiency was not as important a consideration. This would agree with the ultimate goal of national defense, to be able to defend the United States and its allies.

Descriptive Statistics of Chosen Subgroups

The chosen subgroups of Line, Staff, SAC, MAC and TAC have descriptive statistics provided in Table 12. The comparison of these subgroups are important to this study because of the different uses each group has for performance measurement indicators. The comparison is also important to reveal differences because of corporate cultures within the major commands, or line versus staff cultures. Caution should be when interpreting means because much of the data's variability is lost.

Line and Staff Officers Comparisons.

Staff officers had an average mean response (2.7) higher than line officers (2.4), indicating a perception by staff officers that the performance measurements did not respond as favorably to an examination by that group. This may be a result of not having to use the data for immediate management decisions. Staff officers collate the data and may use it for policy decisions but may take very little direct action.

The frequencies of average responses for the subgroups Line, Staff, SAC, MAC, and TAC are provided in Table 13. Line officers feel the measurements fit the criteria to a higher degree than the staff officers. This is reflected by tneir higher frequency (19) of Strongly Agree responses. The staff officers have only (5) of these responses. There is a large discrepancy here that is not shown by the mean responses between line and staff officers. Again this could be a reflection of the corporate culture and how the performance measurements are used at the different levels as stated before.

The line officers also have a greater number of mean Agree responses (177) than the staff officers (140). Th s is similar to the results of the descriptive statistics that show the line officer is more pleased (according to the mean) with certain performance measurements than the staff officers. Again this may reflect the use of performance measurements at different levels.

Based on the Table 13 (frequency distributions), many more of the staff officers chose neutral (78) than the line officer (28) giving more evidence to the finding line officers favor the performance measurements more highly than staff officers. Only one mean response existed in the range of Disagree and that was found in the staff group.

The Wilcoxen Rank Sum Test was used to compare the chosen subgroups of Line, Staff, SAC, MAC, and TAC. The results of these statistical tests are provided in Table 15. The

rationale for the use of this test can be found under this same heading in Chapter III. These subgroups were compared and analyzed for significant differences in group location parameters.

The subgroups Line and Staff were compared and found to have significant differences in their location parameters. The absolute value of the computed z value for this comparison (6.17) was greater than the value of z (alpha=.05)) for a two tailed test (1.64). This indicates the rejection of the null hypothesis that the populations are equal and an acceptance that the location parameters of the populations are significantly different at a 95 percent confidence level. The test also indicates Line is shifted to the left of Staff which has already been shown by descriptive statistics, and an analysis of the frequencies.

Major Command Comparisons.

Between the major commands, MAC responded with the highest mean response at 2.697. This average response is between Agree and Neutral. This command, having an airlift oriented culture, may be dissatisfied with the measurements because they may not reflect the mission of the command.

The second highest mean response at 2.552 belonged to TAC. Again this average response favored the neutral side of Agree. This command consists primarily of fighter aircraft. The crews of fighter aircraft require more general purpose vehicles (sedans, pickups etc . . .) because crew size is limited to two. These general purpose vehicles can be driven

by drivers with less training than is required to drive a bus, normally required by larger crews such as bombers. Trained drivers are usually at a premium in transportation squadrons. If the need for trained bus drivers is reduced then the pressure on transportation support also is reduced. This reduced pressure may be reflected in the response average of the command. The performance measurement indicator may not reflect enough of a contribution to the mission to warrant as much attention to TAC personnel as in SAC.

The lowest mean response at 2.341 belongs to SAC and is the only mean response on the Agree side of the division between Neutral and Agree. This may indeed be a reflection of corporate culture. The SAC nuclear mission makes it very dependent on transportation both for missile and bomber crew vehicles. The control necessary within SAC is a reflection of its nuclear mission. The enormous inherent risk and cost to adequately operationally test this command results in many "simulations" that can only be measured by numbers and not material results. This may necessitate a corporate support structure that must also be under tight control to "fit" the command.

SAC had the largest number of mean Strongly Agree responses (51) by a very large margin (40) over the next highest major command. This would tend to corroborate the findings stated in the descriptive statistics section that SAC favors the performance measurements more than the other

commands. Again the large margin is not adequately reflected in the overall mean but can readily be seen here.

TAC had the largest number of mean Agree responses (177). This agrees with the relative ranking of the overall mean response rate, discussed in the descriptive statistics section, when combined with the Strongly Agree responses and places TAC (188) between SAC (202) and MAC (151) as far as the total agree responses.

The frequency of MAC neutral responses (73) is twice as many as TAC and three times as many as SAC. This ambivalence by MAC may reflect the extent to which the measures are perceived to reflect the mission of the command by the respondents. The relative order of the major commands do not change from the order given by the descriptive statistics.

Using the Wilcoxen Rank Sum Test the subgroups SAC, MAC, and TAC were compared and found to have significant differences in their location parameters. The results of the test are provided in Table 15. The absolute value of the computed z values for these comparisons (7.48,4.86,3.3) for the respective comparisons, SAC vs. MAC, SAC vs. TAC, and MAC vs. TAC, was greater than the value of z (alpha=.05)) for a two tailed test (1.64). This indicates the rejection of the null hypothesis that the populations are equal and an acceptance that the location parameters of the populations are significantly different at a 95 percent confidence level. The test also indicates SAC's location parameters are shifted to the left of MAC and TAC, and TAC's location parameters are

shifted to the left of MAC's which has already been shown by descriptive statistics, and an analysis of the frequencies. The various reasons for the locations and differences between the populations have already been postulated.

Perceptions of Manpower, Budget, Environment, and Actual Cost

Section IV of the survey had the purpose of obtaining the perceptions of the population on the desirability of including manpower, budget, environment, and actual cost figures in their performance measurement indicators. Section IV also measured the populations perceptions on being compared with other base-level transportation organizations based on the performance management indicators in Q13-39.

The dramatic results of Section IV (Q41-45) were difficult to analyze. In Q41, the respondents responded 83 percent in favor, and 12 percent against, manpower being used in the design of performance measurements. Since direct labor was ranked sixth in the top five highest cumulative measurements (see Table 18) and fifteenth in the percentage of agree and strongly agree, the concern over manpower would not be expected to this degree. This would give evidence to a hypothesis that the performance measurements perceived to be the "best" are those given the most attention on base and at higher levels and not those that best reflect the effectiveness and efficiency of the organization.

Over 70 percent responded in favor of environment being considered in performance measurement design (Q42). Since Air

Force Manual 77-310, Volume I provides for the setting of separate standards based on environment, it is surprising that no push has been made to categorize bases according to climate, base mission, surrounding conditions, and adjust performance measurement indicator standards accordingly. What is the case instead is Air Force Manual 77-310, Volume I sets the minimum Air Force standards and the major commands set their minimum levels. The major commands have bases in a large number of different environmental conditions and all are expected to meet the major command goals which vary between the commands.

In Q43, the response of 76 percent in favor and 10 percent against the presence of actual cost figures in dollars in performance measurement indicators was surprising. The fact dollar based performance measurements, such as Average Cost per Mile and Average Cost per Unit, failed to appear in the top ten performance measurements contradicts the result of this question. One possible explanation is, the respondents perceive these particular cost measurements do not adequately reflect the performance of their squadron. Another explanation is, the respondents do not perceive them as being important because the respondents have little control over these figures. Finally, the respondents may feel that these performance measurements should be a part of measurement design but since they are not briefed regularly at the base level, or used to compare bases at the major command level, they are of little importance.

A majority of 61 percent responded in favor and 17.4 percent against, in Q44, the use of squadron budget levels in performance measurement design. This is somewhat contradictory to the results in Table 18 of the relative ranking of the importance of dollar factors as stated above in the analysis of Q43. This result provides insight into why the respondents perceive monetary levels to be important, but the fact that they are not as often briefed or compared across bases may make them seem less important.

Question 45 yielded decisive results with 49 percent against and 22 percent for the use of the performance measurement in Q13-39 being used to compare squadron performance across all bases. Since the comparison of the performance measurements in Q13-39 are regularly used to compare bases within a major command, such as SAC's Peer Competition, these results make it surprising that some effort has not been made to use only those measures that are indeed comparable. A possible reason for these results is these performance measurements do not adequately reflect squadron performance across different variables such as environment, manpower, budget, mission etc . . . Anthony also describes one of the problems of nonprofit organization as being "The difficulty in comparing the productivity among units in the organization." (1:42-44)

The results of the questions on manpower, budget, type of mission, environments, actual cost figures compare favorably with the findings of the literature review. All of these

items have been proposed by researchers as important to performance measurement. The respondents recognized the need for the consideration of these factors.

Summary

The results of the analysis provided evidence to the fact the performance measurements perceived as the "best" are those either measuring customer service, briefed at staff meetings, and used to compare transportation organizations. The conclusions and further recommendations determined from this analysis follow in Chapter VI.

VI. Conclusions and Recommendations

Overview

Performance measurement is an important topic to today's managers because it can be directly translated into dollars. In the case of the Air Force, dollars contribute to combat capability.

The transportation squadron represents a unique problem in the measure of performance. Being a nonprofit organization, the transportation squadron must base its performance on something other than profit. Labor and cost systems are two examples.

The performance measurements currently used by the base-level transportation squadron have only marginally positive support from transportation officers. There exists a substantial faction of the population with negative perceptions of the performance measurements.

This study addressed three major research questions involving the Air Force transportation base-level performance measurement indicators. The conclusions drawn from the analysis of the results of this study will be presented first, and followed by recommendations for possible improvements. Last, suggestions for further areas to be researched, which surfaced as a result of this study, will be given for those researchers wishing to continue along this course of study.

Conclusions

The three research questions addressed in this study are provided below:

1. What are the base-level transportation performance measurement indicators?

2. What are essential characteristics of performance measurement indicators?

3. What are the actual perceptions of transportation officers towards the base-level performance measurements using a reference standard of seven currently accepted criteria?

Research Question 1.

As a result of a review of the literature, key personnel interviews, and analysis of primary data, the base-level transportation performance measurement indicators are provided here:

1. Percentage Hours VDM-Registered
2. Percentage Hours VDP-Registered
3. Percentage Hours VOC-Registered
4. Safety Inspection Overdue-Registered
5. Scheduled Inspection Overdue-Registered
6. Percent Hours VOC-Fire truck
7. Taxi Average Response Time
8. Vehicle Operators Care Rating-Transportation
9. Vehicle Operators Care Rating-Wing
10. Number of Vehicle Abuses

This list of measurement indicators were the top ten measurements most often appearing in both the top five and top ten cumulative ranks, as well as occupying a similar position in the average agree total of all seven performance characteristics.

Research Question 2.

The essential characteristics of performance measurement indicators have been confirmed as those stated by the A.T. Kearney study. The seven characteristics are provided here:

1. Validity
2. Coverage
3. Comparability
4. Completeness
5. Usefulness
6. Compatibility
7. Cost effectiveness

(2:43)

These seven characteristics are confirmed by a review of performance measurement literature. The analysis of the percent response distribution and relative ranking done in this study gave additional evidence to the credibility of these characteristics.

Research Question Three.

Transportation officers surveyed perceive the majority (56.4%) of base-level transportation performance measurement indicators as meeting the seven characteristics of the A.T. Kearney study. There exists however, enough of a perception that the measurements do not meet the criteria (18.9%) and neutral responses (24%), that those measurements receiving a low score should be considered for modification, elimination, or, as a minimum, further evaluation. The specific perceptions of the transportation officers surveyed on each performance measurement indicator and characteristic are provided in Tables 4-10.

The majority of transportation officers surveyed perceive the inclusion of manpower, environment, actual cost figures, and budget in performance measurement design as necessary. The comparison across bases using the current performance measurement indicators is not favorably perceived by the majority of transportation officers surveyed.

Operational Issues.

The perception of performance measurements as a "club" is a result of the failure of the current system to develop reasonable standards of performance. For example, the Air Force minimum standard for VOC is set but each major command can set a more rigorous standard. The Strategic Air Command (SAC) sets a standard more demanding than the Air Force standard by breaking this into 3.7 percent vehicle deadlined for parts rate (VDP) and 6.3 percent vehicle deadlined for maintenance (14:5-3). Another example of an unreasonable performance standard is SAC's goal of "Zero" overdue vehicle scheduled maintenance inspections (14:5-3). On the open ended section of the survey one respondent stated, "Across major command goals such as SAC's Percent Direct Labor goal of 72% by 1990 (a 20% increase) is an impossibility and should not be implemented."

One example of a better operational use of the current performance measurements was suggested,

> Inspector General evaluations should include the evaluations of the current performance measurement indicator levels. Squadrons will then view them as a way to improve performance and not as paperwork." This may result in more interest in transportation performance measurements and more emphasis on developing better up-to-date measurements.

Use of the Vehicle Integrated Management System (VIMS) as a state of the art performance measurement system may not be viewed favorably, as shown by one respondent's comment, "A system should be developed to replace the VIMS system because it is outdated and does not keep up with the changes in budget levels." The system was developed in the 1960's and may not reflect current performance measurement philosophy.

Performance standards need to be designed and interpreted with sensitivity to operational situations. In fact one respondent stated, "Base location, climate, levels of trained personnel, and age of the fleet should all be considered in performance measurement design." The feeling, "Performance measurements should consider 'required' manpower which equals deferred plus authorized positions" was voiced by more than one respondent. Another felt all transportation measures should be "linked to sortie production" because that is the end result of the support function. The "equipment available for deployment (in deployable squadrons)" was the number one priority for one respondent.

Some respondents felt that performance measurement should be tied to customer service. One respondent felt customer service was the "mean time between failure of the vehicles because this represents true availability to the customer and

to the missio Other respondents felt "rejection rate of completed work" and "repetitive maintenance workorders" should be measures of customer ser'ice. A rather well considered suggestion was,

> A customer availability index should be developed that does not factor out things like wrecker run time and contract maintenance and do not establish a standard because many areas are not controllable by transportation." These suggestions concerning customer service, suggest current transportation performance measurements are internally oriented and perhaps need a more external focus.

Recommendations

The practical value of this type of research is the direction it can provide for the transportation community in the future. Sound recommendations must be included in such research that can be of use to the transportation community. In view of the results and conclusions of this study the following recommendations are presented:

1. The inclusion of manpower, environment, actual cost figures, and budget in performance measurement design is necessary and desirable to the majority of transportation officers. Provisions are made for some performance indicators to be adjusted for some of these factors in the regulations but this is not practiced. The provisions should be used for those indicators and extended to those not presently included.

2. If the first recommendation cannot be accomplished, the present performance measurements indicators should not be used to compare transportation organizations across bases. Since the present indicators and the minimum standards are not

adjusted for base function, environment, manpower, these indicators should only be used as an historical comparison in the same organization.

3. If a performance measurement indicator is judged to be one which reflects the performance of the squadron according to criteria such as A.T. Kearney's, then this is the measurement which should be briefed at staff meetings and at major commands. This would help alleviate the situation where the officer is concerned more about the numbers reported than what they represent to the mission. Measurements could be based on an aspect of the "value added" such as customer service. This "value added" could be determined by further research such as a customer audit of using organizations (see Further Research Suggestions section of this chapter).

4. Performance measurements should be modified or redesigned to reflect the benefit to the customer or "value added" and not just internally reflect the state of the organization. Those measurements perceived as reflecting squadron performance and having the greatest attention of the squadron commander were those representing the value to the customer.

5. The open-ended section of the survey asked for any additional thoughts or suggestions on the subject of performance measurements. The respondents comments not presented above are condensed into the recommendations provided here:

- Across the board type measurements are ineffective. Measurements should be people oriented.

- Dollar costs of repairs versus replacement cost should be considered.

- Too much time is spent tracking data when more people are needed in the shops.

- During the winter season track snow removal equipment in the same manner as 463L equipment.

Further Research Suggestions

Commercial application performance measurements and management personnel should be researched as sources of tremendous potential. The wealth of knowledge in the commercial transportation sector can provide new insight into performance measurement design. The commercial sector is more sensitive to "customer service" and "value added" concepts than nonprofit organizations and can only enhance to Air Force transportation unit performance. This study could be extended beyond continental locations, commands other than SAC, MAC, or TAC and services other than the Air Force.

A Delphi technique could be used to gather the experience and knowledge of transportation experts. A Delphi group composed of selected Air Force, Army, commercial, and academic experts could be interviewed. These experts would be asked questions such as, "What are the characteristics of good performance measurement indicators and systems?." and "What is the best way to measure nonprofit organizations?" Once

accomplished, the results would be compared to the results of this study. Recommendations for "new" performance measurement indicators would stem from Delphi technique results and the result of the comparison.

One important aspect of this research should be an external customer audit of organizations using transportation services. The needs of the customer are highlighted in this type of audit. This would allow transportation performance measures to be aligned with customer needs.

Once the commercial sector and the remaining Air Force locations have been studied, the results should be combined and compared. The purpose is to obtain a performance measurement system and indicators that would more accurately reflect the mission of the base-level transportation squadron.

Closure

The great expense of maintaining efficient and effective Air Force transportation support makes good performance measurements vitally important. Without good performance measurements, progress or lack of progress cannot be documented. If progress cannot be measured, the effects of any management changes will not be seen. As managers it is necessary to consistently examine the state of our performance measurements. The more efficiently and effectively the mission can be done, the higher the combat readiness of the Air Force can be.

Appendix A: Selected Productivity Terms

Labor-Hours-"direct or indirect time spent on a task" (34:194)

Partial-Factor Productivity-"the output attributable to a single input factor" (7:20).

Performance-"ratio of actual input to standard output" (34:188)

Productivity-"ratio of real output to real input" (34:188).

Standard Cost-"standard units of input multiplied by the cost per unit" (34:194).

Total-Factor Productivity-"efficiency of transformation of all inputs in combination into outputs" (7:20).

Utilization-"ratio of capacity used to available capacity" (34:188).

Appendix B: Approval Notice

DEPARTMENT OF THE AIR FORCE
HEADQUARTERS AIR FORCE MILITARY PERSONNEL CENTER
RANDOLPH AIR FORCE BASE TX 78150-6001

21 JUN 1989

REPLY TO ATTN OF: DPMYOS

SUBJECT: Request of Survey Approval (Your Ltr, 26 May 89)

TO: AFIT/XP (2d Lt Frisco)
Wright-Patterson AFB 45433-6589

1. After reviewing 1st Lt K. Brewer's "Performance Measurement Survey," we recommend the changes cited in para. 2 need to be incorporated before Air Force personnel may be surveyed using this instrument. Once these changes are made, Survey Control Number (SCN) 89-55 will be assigned to this survey which should appear in the upper right hand corner of the instrument. Authorization will be in effect until 1 November 1989

2. Please advise 1st Lt Brewer to make the following changes:

a. The Privacy Act Statement is not necessary since you are not asking for names or Social Security numbers.

b. The sentence beginning on line 4 of the first page instructions doesn't need the word "defined" and should be revised to refer to sections II and III.

c. Question 3 of the Background Information Section really isn't clear if it's asking for just the time in present position or including previous experience.

d. Your instructions for Section II might be clearer written as "...using the definitions from the Performance Measurement Evaluation Criteria page and the scale provided below."

e. Please add instructions for questions 41-45 of Section IV such as, "Circle the responses which most closely represent your opinions regarding the following questions."

f. For items 41-45, we recommend using "Neither agree nor disagree" rather than "Neutral" as a midpoint for the scales.

3. If there are any questions, please contact Capt Wes Roberts at AUTOVON 487-5680 and he will happily assist 1st Lt Brewer.

Charles H. Hamilton, GM-13
Chief, Personnel Survey Branch

Appendix C: Final Survey

DEPARTMENT OF THE AIR FORCE
AIR FORCE INSTITUTE OF TECHNOLOGY (AU)
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-6583

REPLY TO ATTN OF: LSG (1Lt Kevin Brewer, AUTOVON 785-6569)

SUBJECT: Performance Measurement Survey

TO: Addressee

1. Please take the time to complete the attached questionnaire and return it in the enclosed envelope by 15 July 1989. The questionnaire should take about 35 minutes.

2. This questionnaire assesses how well selected transportation measurements describe unit performance. The data gathered here will become part of an AFIT research project and may influence the way performance measurements are designed in the future. Your responses will be combined with others and will not be attributed to you personally.

3. Your participation is completely voluntary, but we would certainly appreciate your help. For further information, contact Major Robert McCauley at AUTOVON 785-4149.

Robert F. McCauley

Robert F. McCauley, Major, USAF
Instructor Logistics Management

2 Atch
1. Questionnaire
2. Return Envelope

USAF Survey Control No. 89- 55, expires 1 Nov 89

Please tear off this sheet to use while answering the questionnaire.

Performance Measurement Evaluation Criteria

a. Validity--Does the measure track real productivity? The measure should change to accurately reflect changes in real productivity.

b. Coverage--Does the measure capture all uses of a resource?

c. Comparability--Is it possible to compare the measure with the same measure at another location or at a different time? For example, does the measure mean the same thing it did last year or the same thing from one base to another?

d. Completeness--Does the measure account for important resources? A measure needs to track the resources that matter most in the operation. Labor and cost are two examples of important resources.

e. Usefulness--The measure must help lead the manager to an effective decision or action.

f. Compatibility--Is the measurement compatible with the information systems or data systems currently used?

g. Cost Effectiveness--Do the benefits derived from collecting this information exceed the cost of gathering, storing, and analyzing the measure?

PRIVACY ACT STATEMENT

In accordance with paragraph 8, AFR 12-35, the following information is provided as required by the Privacy Act of 1974:

a. Authority

(1) 5 U.S.C. 301, Departmental Regulations; and/or

(2) 10 U.S.C. 8012, Secretary of the Air Force, Powers, Duties, Delegation by Compensation; and/or

(3) DOD Instruction 1100.13, 17 Apr 68, Surveys of Department of Defense Personnel; and/or

(4) AFR 30-23, 22 Sep 76, Air Force Personnel Survey Program.

b. Principal Purposes. The survey is being conducted to collect information to be used in research aimed at illuminating and providing inputs to the solution of problems of interest to the Air Force and/or DOD.

c. Routine Uses. The survey data will be converted to information for use in research of management related problems. Results of the research, based on the data provided, will be included in written master's theses and may also be included in published articles, reports, or texts. Distribution of the results of the research, based on the survey data, whether in written form or presented orally, will be unlimited.

d. Participation in this survey is entirely voluntary.

e. No adverse action of any kind may be taken against any individual who elects not to participate in any or all of this survey.

Performance Measurement Survey

The purpose of this study is to measure your perceptions of selected transportation performance measurements, based on the seven criteria for effective productivity measures developed by A.T. Kearney, Inc. for the National Council of Physical Management. The definitions for the seven criteria are on the previous page and should be used to answer questions in sections II and III of the survey. It is important to obtain your honest opinion, there is no "right" answer.

I. Background Information

1. Your rank _____

2. Primary AFSC _____

3. Number of years in present type position (include previous) command or analytical transportation positions):_____

4. Number of years in transportation or related field: __

5. Number of personnel in your squadron: _____

6. Are you a line or staff officer? Line__ Staff__

7. Your age ____

8. What is the highest level of education that you have completed?

   ___ Completed college. Major_____
   ___ Some graduate work
   ___ A graduate degree. Degree_____
   Major_____

9. Current Major Command _____

10. Base function: Bomb Wing ___ ___ Fighter Wing
    Missile Wing ___ ___ Dual Wing.
    Which two?_____
    Other _____

11. Name of base: _____

12. Present job title: _____

II. Evaluation of Vehicle Integrated Management System Performance Indicators

Please mark your perception of the following performance measurements at your present base or position using the definitions from the Performance Measurement Evaluation Criteria page and the scale provided below.

Strongly agree Agree Neutral Disagree Strongly disagree

1---------2-------3-------4-----------5

An answer of strongly disagree would mean you felt that this measure did not meet the criteria at all. Neutral means you neither agree nor disagree. An answer of strongly agree would mean the measure meets the criterion completely.

13. Average cost per unit - Registered

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

14. Safety inspection overdue - Registered

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

15. Scheduled inspections overdue - Registered

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

16. Number of work orders opened

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

17. Hours delay code C

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

18. Percent direct labor

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

19. Percent indirect labor productive

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

20. Percent indirect labor non-productive

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

21. Percent VDM - Registered

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

22. Percent VDP - Registered

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

23. Percent VOC - Registered

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

24. Percent hours VOC - B & K

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

25. Percent hours VOC - D & M

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

26. Percent hours VOC - W

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

27. Percent hours VOC - E (463L)

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

28. Percent hours VOC - E (MHE)

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

29. Percent hours VOC - Fire truck

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

30. Percent hours VOC - Refueling

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

31. Percent hours VOC - C & L

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

32. Percent hours VOC - Non-registered

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

33. VDM/Direct labor percent - Registered

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

34. Average cost mile - Admin fleet

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

III. Vehicle Operations Performance Indicators

Strongly agree Agree Neutral Disagree Strongly disagree
1----------2-------3-------4----------5

35. Taxi - average response time

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

36. Vehicle operator care rating - wing

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

37. Vehicle operator care rating - transportation squadron

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

38. Number vehicle misuses

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

39. Number of vehicle abuses

a. Validity	1	2	3	4	5
b. Coverage	1	2	3	4	5
c. Comparability	1	2	3	4	5
d. Completeness	1	2	3	4	5
e. Usefulness	1	2	3	4	5
f. Compatibility	1	2	3	4	5
g. Cost Effectiveness	1	2	3	4	5

IV. General

40. Please list ten of the above performance measurements by number and in order of importance from most to least important that you feel are important to you in managing your squadron.

___ ___ ___ ___ ___ ___ ___ ___ ___ ___

Most Least

Circle the responses which most closely represent your opinions regarding the following questions.

41. Authorized versus assigned manpower levels should be considered in performance measurement design.

    1. Strongly agree
    2. Agree
    3. Neither agree nor disagree
    4. Disagree
    5. Strongly disagree

42. Environmental conditions (such as ambient temperature) should be considered in performance measurement design.

    1. Strongly agree
    2. Agree
    3. Neutral (Neither agree nor disagree)
    4. Disagree
    5. Strongly disagree

43. Actual cost figures (in dollars) should be present in performance measurement indicators, where possible.

    1. Strongly agree
    2. Agree
    3. Neutral (Neither agree nor disagree)
    4. Disagree
    5. Strongly disagree

44. Squadron budget levels should be considered in performance measurement design.

    1. Strongly agree
    2. Agree
    3. Neutral (Neither agree nor disagree)
    4. Disagree
    5. Strongly disagree

45. The performance measurements in questions 13-39 should be used to compare squadron performance across all bases.

    1. Strongly agree
    2. Agree
    3. Neutral (Neither agree nor disagree)
    4. Disagree
    5. Strongly disagree

If you have any additional performance measurements that you use in managing your squadron or additional thoughts on this subject please write them below and on the back of the survey. When finished place the survey into the envelope provided and mail. Thank you for your cooperation; your help is appreciated.

Appendix D: Reminder Card

Dear Transporter:
Recently you received a survey called
"Performance Measurement Survey".
Because only 67 transporters were surveyed, your response is very important. In fact, your response is essential to me being able to finish my thesis. Please send your response. If you already have, I thank-you.

1Lt Kevin Brewer
Transportation Officer

Appendix E: Histogram of Subgroups

LINE

```
  LOW     HIGH
       N
-------------
  1.6     1.9  *********
    10         *********
  1.9     2.2  ****************************************************
    58         ****************************************************
  2.2     2.5  ****************************************************
    57         ****************************************************
  2.5     2.8  ************************************
    44         ************************************
  2.8     3.1  ********************************
    39         ********************************
  3.1     3.4  ********
     9         ********
  3.4     3.7  ******
     7         ******

CASES INCLUDED 224    MISSING CASES 0
```

STAFF

```
  LOW     HIGH
       N
-------------
  1.5     1.9  **
     2         **
  1.9     2.3  ******************************
    33         ******************************
  2.3     2.7  *********************************************
    57         *********************************************
  2.7     3.1  ****************************************************
    63         ****************************************************
  3.1     3.5  ************************************
    46         ************************************
  3.5     3.9  ***********
    14         ***********
  3.9     4.3  **
     3         **

CASES INCLUDED 224    MISSING CASES 0
```

SAC

```
  LOW     HIGH
        N

 ------------
   1.4     1.7  ******
        6       ******
   1.7     2.0  ********************************************
       45       ********************************************
   2.0     2.3  **************************************************
       54       **************************************************
   2.3     2.6  ********************************************
       46       ********************************************
   2.6     2.9  ********************************
       33       ********************************
   2.9     3.2  ********************************
       36       ********************************
   3.2     3.5  ****
        4       ****
```

CASES INCLUDED 224 MISSING CASES 0

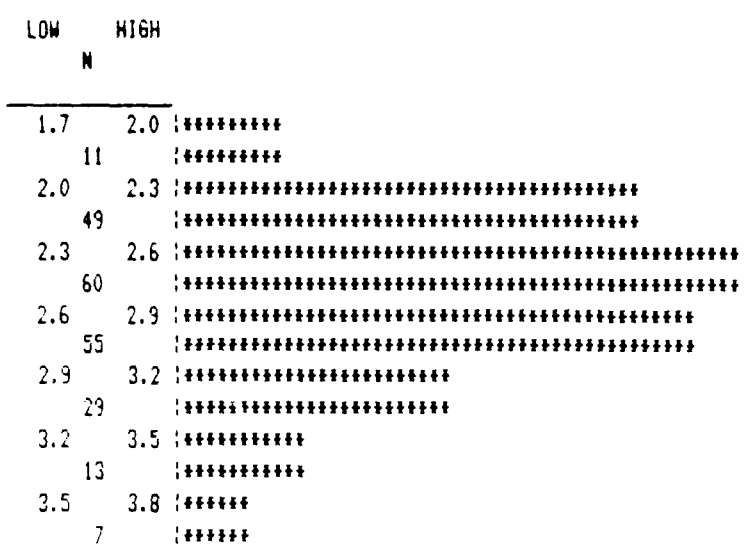
MAC

```
  LOW     HIGH
        N

 ------------
   1.2     1.6  *
        1       *
   1.6     2.0  *****
        6       *****
   2.0     2.4  ************************************************
       63       ************************************************
   2.4     2.8  ********************************************
       56       ********************************************
   2.8     3.2  **************************************************
       64       **************************************************
   3.2     3.6  ***************
       13       ***************
   3.6     4.0  ************
       15       ************
```

CASES INCLUDED 224 MISSING CASES 0

TAC



CASES INCLUDED 224 MISSING CASES 0

Bibliography

1. Anthony, Robert N. and David W. Young. Management Control in Nonprofit Organizations. Homewood IL: Richard D. Irwin, Inc., 1984.

2. Astin, Jared A. and Christopher D. Ruff. Determining Managerial Methods of Productivity Measurement within Civil Engineering Design Units. MS Thesis. AFIT/GEM/LSM/84S-1. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1984 (AD-A146934).

3. Baker, B.O. and others. "Weak Measurements vs. Strong Statistics: An Empirical Critique of S.S. Stevens' Proscriptions on Statistics," Educational and Psychological Measurement, 26: 291-309, 1966.

4. Ballou, Ronald H. "Measuring Physical Distribution Performance," The National Council of Physical Distribution Management Annual Meeting 1983: 836-841. The National Council of Physical Distribution Management, 1983.

5. Baumgartel, Captain Gary P. and Captain F. Fink. Productivity Measurement in a Base Level USAF Civil Engineering Organization. MS Thesis. AFIT/GEM/LSM/84S-1. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, May 1984 (AD-A146934).

6. Bras, Colonel Victor. "Contractor-Operated Parts Stores: Is Change Overdue?" Air Force Journal of Logistics: 12-17 (Fall 1988).

7. Campbell, John P. and R.J. Campbell. Productivity in Organizations. San Francisco: Jossey-Bass Publishers, 1988.

8. Conlon, Patrick J. and William A. Townsend. "The "How To" of Transportation Productivity Programs," The National Council of Physical Distribution Management Annual Meeting 1984: 53-74. The National Council of Physical Distribution Management, 1984.

9. Coyle, John J. and others. Transportation. New York: West Publishing Company, 1986.

10. Crawford, Karlene M. An Analysis of Performance Measurement Systems in Selected Just-in-Time Operations. PhD dissertation. University of Georgia, Atlanta GA, 1988.

11. Deakin, B.M. and T. Seward. Productivity in Transport. Cambridge: Cambridge University Press, 1969.

12. Demers, W.A. "Grading Contractor Performance," Military Logistics Forum,4: 38-42 (May 1988).

13. Department of the Air Force. Command, Air Force Vehicle Integrated Management System (CAFVIMS). AFM 77-310, Volume IV. Washington: HQ USAF, 9 November 1979.

14. -----. Management Guide for Resource Managers. SACP 11-16. Offut Air Force Base: HQ SAC, 15 May 1986.

15. -----. Management Guide for Transportation Commanders. SACP 75-5. Offut Air Force Base: HQ SAC, 21 December 1987.

16. -----. Transportation Squadron Commanders Guide. 15AF 75-1. March Air Force Base: HQ 15AF, 14 July 1986.

17. -----. Vehicle Maintenance Management. AFM 77-310, Volume II. Washington: HQ USAF, 15 March 1978.

18. -----. Vehicle Operations Management. AFM 77-310, Volume I. Washington: HQ USAF, 22 May 1988.

19. Department of Defense. DOD Productivity Program. DOD Directive 5010.31. Washington: Government Printing Office, 27 April 1979.

20. -----. Productivity Enhancement Measurement, and Evaluation-Operating Guidelines and Reporting Instructions. DOD Instruction 5010.34. Washington: Government Printing Office, 4 August 1975.

21. Dillman, Don A. Mail and Telephone Surveys. New York: John Wiley and Sons, 1978.

22. Emory, C. Williams. Business Research Methods. Homewood, Illinois: Irwin Inc., 1985.

23. Fallon, William K., American Management Handbook. New York: American Management Association Inc., 1983.

24. Fogerty, Donald W., and Hoffman, Thomas R. Production and Inventory Management. Cincinatti: South-Western Publishing Co., 1983.

25. Gautschi, David A. Productivity and Efficiency in Distribution Systems. New York: Elsevier Science Publishing Co., Inc., 1983.

26. Goldratt, Eliyaker and Robert E. Fox. "The Fundamental Measurements," The Theory of Constraints Journal: August/September 1988.

27. Greenberg, Leon. A Practical Guide to Productivity Measurement. Bureau of National Affairs, Inc., 1973.

28. Greene, James H. Operations Management: Productivity and Profit. Reston, VA: Reston Publishing Company, 1984.

29. Hanley, Lieutenant Colonel William F. and Captain Frederick C. Smith. USAF Civil Engineering Work Force Productivity: An Analysis of Non-Behavioral Factors Affecting Planning Technician Labor Manhour Estimates for In-House Work Requirements. MS Thesis. AFIT/GEM/LSM/76S-1. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, May 1976 (AD-A030244).

30. Hansen, General Alfred G. "The AFLC Quality Program," Journal of Business Logistics, Council of Logistics Management 10: 3-9 (Spring 1989).

31. Howell, Captain R.C. and Lieutenant J.D. Van Sickle. Perceptions of a Methodology for the Development of Productivity Indicators. MS Thesis. AFIT/GEM/LSM/82S-1. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1982 (AD-A123026).

32. Johnson, Major Jont, Chief, Transportation Squadron. Personal interview. Wright-Patterson AFB OH, June 5, 1989.

33. Kaneda, Captain Kenneth G. and Captain Robert M. Wallett. Development of Productivity Measures For the Design Section of a Base Level Civil Engineering Organizations. MS Thesis. AFIT/GEM/LSM/80S-1. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, 1980 (AD-A087496).

34. Kearney, A.T., Inc. Measuring and Improving Productivity in Physical Distribution. Oak Brook, IL: A.T. Kearney, Inc, 1984.

35. King, Lieutenant Colonel Dan E., Transportation Squadron Commander. Personal interviews. 87TH Transportation Squadron, Minot AFB ND, 10 June 1989.

36. Lowry, Ralph W. Measuring Productivity of Depot-Level Aircraft Maintenance in the Air Force Logistics Command. MS Thesis. AFIT/GEM/LSM/82S-1. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1982 (AD-A123026).

37. Novak, Frederick V. "MDC Data Integrity," Logistics Management Review, SACRP 66-3, 12: 2 (December 1987).

38. Ott, Lyman. An Introduction to Statistical Methods and Data Analysis. Boston, MA: PWS-Kent Publishing Company, 1984.

39. Schmenner, Roger W. Production/Operations Management. Chicago: Science Research Associates, Inc., 1987.

40. Searock, Charles J. Jr. "Enhanced Mission Support--The Focus of Base Supply," :Logistics Management Review, SACRP 66-3, 12: 1 (December 1987).

41. Sink, Scott D. "Performance and Productivity Measurement: The Art of Creative Scoreboards." Industrial Engineering: 86-90, (February 1986).

42. Statistix II. Product Catalog. NH Analytical Software, Roseville, MN, undated.

43. Stephenson, Fred J. Transportation U.S.A. New York: Addison-Wesley Publishing Company, 1987.

44. Temple, Barker, and Sloane Inc. Transportation Strategies for the Eighties. Oak Brook, IL: National Council of Physical Distribution Management, 1982.

45. Tuttle, Thomas C. "Productivity, Measurement Methods: Classification, Critique, and Implications for the Air Force." Unpublished report. Maryland Center for Productivity and Quality of Working Life, College Park MD., September 1981.

46. Tyworth, John E. and others. Traffic Management. Reading, MA: Addison-Wesley Publishing Company, 1987.

47. Williams, Ben. "Durability, Quality of Goods Should Measure Productivity," Dayton Daily News: May 15, 1989.

Vita

First Lieutenant Kevin N. Brewer [redacted] attended Western Illinois University, from which he received a Bachelor of Science in Biology in December 1981. In 1986, he was commissioned into the Air Force. He served as Vehicle Operations Officer and Plans and Programs Officer at Minot Air Force Base until entering the School of Logistics, Air Force Institute of Technology, in June 1988.

[redacted]

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFIT/GLM/LSM/89S-3	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION School of Systems and Logistics 6b. OFFICE SYMBOL (*If applicable*) AFIT/LSM	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (*City, State, and ZIP Code*)	7b. ADDRESS (*City, State, and ZIP Code*)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (*If applicable*)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (*City, State, and ZIP Code*) Air Force Institute of Technology Wright-Patterson AFB OH 45433-6583	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO / WORK UNIT ACCESSION NO.

11. TITLE (*Include Security Classification*)
PERCEPTIONS OF AIR FORCE BASE-LEVEL TRANSPORTATION OFFICERS TOWARDS THE EFFECTIVENESS OF AIR FORCE BASE-LEVEL TRANSPORTATION PERFORMANCE MEASUREMENT INDICATORS

12. PERSONAL AUTHOR(S)
Kevin N. Brewer, M.S., 1st Lt, USAF

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (*Year, Month, Day*)	15. PAGE COUNT
MS Thesis	FROM _____ TO _____	1989 September	131

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (*Continue on reverse if necessary and identify by block number*)
FIELD	GROUP	SUB-GROUP	Productivity, Performance, Performance, Transportation
13	06		
06	01		

19. ABSTRACT (*Continue on reverse if necessary and identify by block number*)

Thesis Advisor: Major Robert McCauley
Instructor of Logistics Management
Department Of Logistics Management

Approved for public release: IAW AFR 190-1.

LARRY W. EMMELHAINZ, Lt Col, USAF 14 Oct 89
Director of Research and Consultation
Air Force Institute of Technology (AU)
Wright-Patterson AFB OH 45433-6583

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Robert McCauley, Maj, USAF	22b TELEPHONE (*Include Area Code*) (513)255-4149 / 22c. OFFICE SYMBOL LSM

DD Form 1473, JUN 86 *Previous editions are obsolete.*

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

The purpose of this study was to examine the perceptions of selected base-level transportation performance measurement indicators. These included the Vehicle Integrated Management System (VIMS), as well as certain vehicle operations indicators. Transportation squadron commanders and certain staff officers within the continental United States belonging to the Strategic Air Command (SAC), Tactical Air Command (TAC), and Military Airlift Command (MAC) were surveyed. The study examined the officers' perceptions against seven criteria for good performance measurements developed by the A.T. Kearney consultant firm.

1. Validity
2. Coverage
3. Comparability
4. Completeness
5. Usefulness
6. Compatibility
7. Cost Effectiveness

An examination of the perceptual surveys found that SAC personnel have a slightly more favorable perception that the performance measurements met the criteria than the other commands, followed by TAC, and then MAC. A comparison between Line and Staff officers revealed Line officers have a slightly more favorable perception that the performance measurements met the criteria.

Analysis of the rankings of the performance measurements indicated the five most important to the respondents were

1. Percentage Hours VDM-Registered
2. Percentage Hours VDP-Registered
3. Percentage Hours VOC-Registered
4. Safety Inspection Overdue-Registered
5. Scheduled Inspection Overdue-Registered

One of the recommendations provided to continue the study based on a Delphi approach using other military service transportation personnel, members of the academic transportation community, and commercial transportation personnel.